

34 spArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
 101 ATGATATGCAAACTTGAGCAATTCCTTTACTAACGACTTACCGGT 150
 51 ValAspTyrThrTyrAlaGlnTyrLeuGluLysSerCysCysLeuGly 67
 151 GTGATTATATGCGACCGCAATCTTAGAAAAATCATGCTGCTAGCAGA 200
 67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
 201 GGCTTAAAGAAATTATGGTTGGTGTCTGATGGAAGAAATTCGGTATGCA 250
 84 erGluAsnGlySgluGluPhePheIleProValLeuAlaGlyLeuPhe 100
 251 GTGAAACCTGTAAGAAATCTTATCTCTGATATAGCCGGTTATATTATA 300
 101 GlyValGlyValAlaProThrAsnGlnIleTyrThrLeuArgLysLeuVal 117
 301 GGTCGCGGTGCTGCCAATGATGATTTACCTGCTAGCGATGATGGT 350
 117 LHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysGly 134
 351 TCACAGATTAGGCATCTTACCAACCAATGTGATTTAGTTCTAAAAAG 400
 134 LysLeuAspLysValIleThrValGlnLysThrValIleAlaIleLysThr 150
 401 GATTGATTAAGATTAACTGATACAAAAAAGCGTAACGCTATTAAACC 450
 151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAs 167
 451 ATTTGATATTGACAGCAAGTGAATATAGAGTTATCAATTCATGCA 500
 167 pAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPhe 184
 501 CAACCTTATTAATAAAACACTCCACAAAGGTTTCAAGATCAAGTATGTA 550
 184 ySThrValGlyValAsnArgLysGlnValAlaLeuIleMetAsnSer 200
 551 AAACGTGAGAAATTAAACCGCAAGAACAAAGTTGCTTATTAATGAACT 600
 201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGlyAsnIle 217
 601 TCGGCTTCAACCGGTTTGCAGAAAGTGCGCAACTTACTCATGAAATTT 650
 217 uValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
 651 GGTCACTGATTTTCTCACGCTAGAGATCCAATTTATGTGAACCAAGTTT 700
 234 erProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
 701 CACGAGGACGCGTATTTTAACTGATGATCAATTCATCATGATGTTTTGCT 750
 251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetIle 267
 751 ATGTTTACTACTTTAGCGCTATCTAACTGTGTTTCTGATTTGCATGCTT 800
 267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrIleS 284
 801 AACGAAATTTGACAGAGACTTTTAAAAAACCTGCAAGATTACAAAT 850
 284 ySSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
 851 GTTCAAGGCTTATCTTGTACCGACTTTGTTGCAATTTCTTAATGAAAGT 900
 301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGly 317
 901 GAATTACTCGATAATAATGATTTATCAAAATTTAGTGAATTTGATCTGG 950
 317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
 951 CGAGACCTCTTATCTAAAGAAATTTGGAGCTGTGCTAGAGCTTTTA 1000

334 snLeuProGlyValAlaArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
 1001 ATTTACCGGGTGTCTGTCAGAGCTATGGTTTAAACAGAAACACTCTGCA 1050
 351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVal 367
 1051 ATTATTATCACACCGAAGGCGATGATTAACCCAGCTCTCTGGCAAGT 1100
 367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
 1101 TGTGCCATTATTTAAGCAAAAGTTATCGATCTTGTGATACATAAAAACTT 1150
 384 euGlyProAsnArgArgGlyGlyValCysValLysGlyProMetLeuMet 400
 1151 TGGGCCCGAAGACAGCGTGGAAGTTGTGTAAAGGCTCCTATGCTTATG 1200
 401 LysGlyTyrValAlaAspAsnProGluAlaThrArgGluIleIleAspGlu 417
 1201 AAAGTTATGATGATTAATCCAGAAACCAAGAAATCATAGATGAGA 1250
 417 uGlyTrpLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
 1251 AGTTGGTTCACACACAGAGATTTGGGTATTTACGATGAAGAAAAACATT 1300
 434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
 1301 TCTTTATCGTGATCGTTTGAAGCTTTATCAAAATCAAAAGATATCA 1350
 451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
 1351 GTACCACTCTCTGAATTAAGAAATCTGTTTTCACATCCAAATATTTT 1400
 467 eaSPAlaGlyValAlaGlyValProAspProIleAlaGlyLeuProG 484
 1401 TGATGCCGGGTTGTCGCGCTTCACAGATCTTATAGCTGGAGCTTCGG 1450
 484 lYalAlaValValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
 1451 GAGCTGTTGTTGTAATTAAAGAAAGAAATCTATGACTGAAAGAACTA 1500
 501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuAlaGlyL 517
 1501 ATGGATTACCTTCTGATCAAGTTTCAAAATGCAAAAGCTTTCGCGTGG 1550
 517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
 1551 TGTCCGTTTGTGACGAGAGTACTTAAGGTCTCACTGGTAAATTTGACG 1600
 534 lYysAlaIleLeuArgGluIleLeuLysLysProValAlaLysMet 548
 1601 GTAAGCAATTAGAGAAATTAAGTAAACCAAGTTGCTTAAGATG 1644
 seq_name: /SID82/gcdata/geneseq/geneseq/NM1999.DAT:AAx8396
 seq_documentation_block:
 ID AAx84396 standard; DNA; 1644 BP.
 XX
 AC AAx84396;
 XX
 DT 09-SEP-1999 (first entry)
 XX
 DE L. lateralis luciferase coding sequence.
 XX
 KW Luciferase; surfactant-tolerant; firefly; intracellular ATP assay; ss.
 OS Luciola lateralis.
 XX
 PN W09933997-A1.
 XX
 PD 08-JUL-1999.
 XX
 PF 24-DEC-1998; 98WO-JP05864.
 XX

PR 26-DEC-1997; 97JP-0361022.
 PA (KIRK) KIRKMAN CORP.
 XX Hattori N, Murakami S;
 XX WPI, 1999-419109/35.
 DR P-PSDB; AAY22183.
 PT Modified luciferase tolerant to surfactants and useful for assay of
 PT Intracellular ATP
 XX
 PS Example 2; Page 43-48; 56pp; English.
 XX This sequence encodes a luciferase lateralis luciferase of the invention.
 CC The protein is a surfactant-tolerant luciferase which is derived from
 CC firefly luciferase by replacement of the glutamic acid residue at
 CC position 490 by another amino acid. The luciferase can be used in an
 CC assay of intracellular ATP in biological samples containing cells, for
 CC example for establishing the presence of foodstuffs, drugs,
 CC agrochemicals, beverages and environmental samples. The activity of the
 CC surfactant-tolerant luciferase is not lowered by the presence of high
 CC concentrations of surfactant and therefore an assay of high accuracy can
 CC be carried out in a simple manner.
 CC
 XX Sequence 1644 BP; 529 A; 262 C; 348 G; 505 T; 0 other:

alignment_scores:
 Quality: 2821.00 Length: 548
 Ratio: 5.148 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.818

alignment_block:
 US-09-581-241-4 x AAY84396 ..

Align seg 1/1 to: AAY84396 from: 1 to: 1644

```

1 MetGluAsnMetGluAsnAspGluAsnIleValTYRGLYProGluProph 17
1 ATGGAAACATGAGACGATGAAATATTGTGTATGCTGATGCAACCAT 50
17 eTYrProIleGluGlySerAlaGlyAlaGluLeuArgLysTYRMeTA 34
51 TTACCCATTGACAGAGGATCTGCTGGAGCACAATTGCGAGATATATG 100
34 sPARgTYrAlaLysLeuGlyAlaIleAlaIlePheThrAsnAlaLeuThrgly 50
101 ATCGATATGCAAAACCTTGAGCAATTCCTTTACTAACGCACCTTACCGGT 150
51 ValAspTYrThrTYrAlaGluTYrLeuGluLysSerCysGlyLeuGly 67
151 GTGGATTATACGTACGCCGAATCTAGAAAAATCATGCTGTAGAGAGA 200
67 uAlaLeuLysAsnTYrGlyLeuValValAspGlyArgIleAlaLeuGly 84
201 GGCTTTAAAGAAATATGGTTGGTTGTTGATGAGAAATTCGCTATGCA 250
84 erGluAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
251 GTGAAAACGTGCAAAATCTTTATTCCTGATTTAGCCGTTTATTTATA 300
101 GlyAlaGlyValAlaProThrAsnGluIleTYrThrLeuArgGluLeu 117
301 GGTTGCTGGTGGCTCCAACTAATGATTTACCTACCTACCTGATGATGCT 350
117 lHisSerLeuGlyIleSerLysProThrIleValAlaPheSerSerLys 134
351 TCACAGTTTAGGCATCTCTAAGCAACAAATGATTTAGTTCTCAAAAAAG 400
134 lYleuAspLysValIleThrValGluLysThrValThrAlaIleLysThr 150

```

```

401 GATTAGATAAAGTTATTAAGTACAAAAACGTAACGTAATTAAC 450
151 lIleValIleLeuAspSerLysValAspTYrArgLysTYrGlnSerMeTAS 167
451 ATTGTTATATTGACACACAAAGGATTAAGGTTATCAATTCATGGA 500
167 pAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheI 184
501 CAACCTTTATTAATAAAAAACACTCCACAAAGGTTCCAAAGATCAAGTTTA 550
184 ySThrValGluValAsnArgLysGluValAlaLeuIleMetAsnSer 200
551 AAACGTGAAAGTTAAACGCAAGAACAGAGTCTTATTAATGAACCTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGluLeuThrHisGluAsn 217
601 TCGGGTTCAACCGGTTTCCAAAGGTGTGCACTTACTCATGAAAAATAT 650
217 uValThrArgPheSerHisAlaArgAspProIleTYrGlyAsnGlnValS 234
651 CGTCACTAGATTTTCTCAGCGTAGAGATCCAAATTTATGGAACCAAGTTT 700
234 erProGlyThrAlaIleLeuThrValAlaProPheHisGlyPheGly 250
701 CACGAGCAGCGCTATTTTAACTGATGATACCTTCATCATGATGTTGGT 750
251 MePheThrThrLeuGlyTYrLeuThrCysGlyPheArgIleValMetIle 267
751 ATGTTTACTACTTAGGCTATTAACCTTGCTTTCGATTTGATATCTT 800
267 uThrLysPheAspGluLysThrPheLeuLysThrLeuGlnAspTYrLysC 284
801 AACGAAATTTGACGAAGACCTTTTAAACACCTGCAAGATTTCAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTTCAAGCGTTATTTCTGACCGACTTGTTCCTTCAATCTTAAAGAACT 900
301 GluLeuLeuAspLysTYrAspLysSerAsnLeuValGluIleAlaSerG 317
901 GAATTAATCGATTAATAATATGATTTATCAAAATTTAGTCAAAATTC 950
317 yGlyAlaProLeuSerLysGluIleGlyAlaValAlaArgArgPheA 334
951 CCGAGACCTTTATTAAGAAATGGAAGCTGTGCTAGACGTTTAA 1000
334 snLeuProGlyValArgGlnGlyTYrGlyLeuThrGluThrSerAla 350
1001 ATTTACCGGGGTTCGTCAGAGCTATGCTTTAACGAAACACCTTGCA 1050
351 lIleIleIleThrProGluLysAspLysProGlyAlaSerGlyLysVa 367
1051 ATTATATATCACACCGAAGCATATTAACACGATGCTTGTGCAAAAGT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThr 384
1101 TGTCGCATTTATTAAGCAAAAGTTATTCATCTTCACTTAAAAAACTT 1150
384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
1151 TGGGCCGCAACAGACGTCGAGAAAGTTTGTGTAAGGTCCTATGCTTATG 1200
401 LysGlyTYrValAspAsnProGluAlaThrArgGluIleIleAspGlu 417
1201 AAAGCTTATGTAGATTAATCCAGAGCAACAGAAATCATATGATGAGA 1250
417 uGlyThrLeuHisThrGlyAspIleGlyTYrTYrAspGluGluLysHisP 434
1251 AGGTGCTTGCACACAGAGATATGCGTATTCAGATGAAGAAACATTT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTYrLysGlyTYrGln 450
1301 TCTTTATCTGATGCTTGAAGTCTTTAATCAATCAAAAGATATCAAA 1350

```

```

451 ValProAlaGluLeuGluSerValLeuGluHisProAsnIleph 467
1351 GTACACACCTGTAATAGAAATCTCTTTTGCACATCCAAATATATTTT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyLeuLeuProG 484
1401 TGATCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1450
484 lAlaValAlaValLeuLeuGlyGlySerMetThrGlyGlyVal 500
1451 GAGCGTGTGTGCTTAAAGAAAGAAATCTAGCTGAAAAAGAGTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgIyl 517
1501 ATGGATTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
1551 TGTCCGTTTGTGAGCAGAGTACCTAAAGCTCTCAGCTGTAATAATTGACG 1600
534 lLysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
1601 GTAAAGCAATTAGAGAAATAGTCAAGAAACAGTTGCTAAGATG 1644

```

seq_name: /SIDS2/gcdata/geneseq/geneseq/NA1999.DAT:AAK25718
 seq_documentation_block:
 ID AAK25718 standard; cDNA to mRNA; 1644 BP.

```

AC AAK25718;
XX
XX 21-MAY-1999 (first entry)
DE Firefly Luciferase gene #4.
XX
XX Bioluminescent protein; catalytic efficiency; stability; firefly;
KW Luciferase; chimeric; luciola cruciata; luciola lateralis;
KW Photinus pyralis; primer; PCR; amplification; ss.
XX
XX Luciola lateralis.
OS Synthetic.
XX
XX W09902697-A1.
PN
PD 21-JAN-1999.
XX
XX 30-JUN-1998; 98MO-JP02936.
PF
XX 08-JUL-1997; 97US-0051917.
PR
XX (KIKK) KIKKOMAN CORP.
PA
XX Hirokawa K, Kajiyama N, Murakami S.
PI WPI, 1999-120898/10.
XX P-PSDB; AAM99367.
DR
XX
XX New bioluminescent protein with improved properties - has greater
PT catalytic efficiency and stability and is obtained by modification
PI of natural precursors
XX
XX Example 5; Page 35-36; 53pp; Japanese.
PS
XX The invention relates to the generation of bioluminescent proteins with
CC improved catalytic efficiency and stability. The proteins are generated
CC by addition, deletion or substitution of amino acids from a precursor
CC protein e.g. firefly luciferase, or by constructing chimeric luciferase
CC proteins from luciferase genes from e.g. luciola cruciata, L. lateralis
CC and Photinus pyralis. This sequence represents an example of a mutated
CC luciferase gene of the invention.
XX
XX Sequence 1644 BP; 528 A; 260 C; 350 G; 506 T; 0 other.
SQ

```

alignment_scores:
 Quality: 2817.00 Length: 548
 Ratio: 5.150 Gaps: 0
 Percent Similarity: 99.818 Percent Identity: 99.818

alignment_block:
 US-09-581-241-4 x AAK25718 ..

Align seg 1/1 to: AAK25718 from: 1 to: 1644

```

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProph 17
1 ATGGAAGAAACATGGAGAACGATGAAATATGTTGATGCTGCTGAACCAT 50
17 eTyrProIleGluGluGlySerAlaGlyAlaGluLeuArgLysTyrMet 34
51 TTACCCCTATTGAAGAGGATCTGCTGAGCACAATTGCCAAGATATATG 100
34 sParGlyAlaLysLeuGlyValAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 ATGATATATGAAACCTTGAGAGCAATTGCTTTTACTTACGACTTACCGGT 150
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGly 67
151 GTGATATATACGTACGCGCAATACCTTACAAAAATCATCTGCTTACGAGA 200
67 uAlaLeuLysAsnTyrGlyLeuValAlaAspGlyArgIleAlaLeuCys 84
201 GCCTTAAAGAAATTTGCTTTGCTTTGATGAGAAATGCGCTTATATGCA 250
84 eArgLysCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
251 GTGAAGACTGTGAAGAAATCTTATCTCTGATATAGCGGTTTATTTATTA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeu 117
301 GGTGCGGTGTGCTCCACATAAGATTTACACTGCTACGTAATTTGCT 350
117 lHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysG 134
351 TCACAGTTTAGCATCTCTTAAGCAACAATTGATTTAGTTCTTAAAGAA 400
134 lLeuAspLysValIleThrValGluLysThrValThrAlaIleLysThr 150
401 GATTAGATAAAGTTATACGTACAAAAACGGTAACTGCTATTAAACC 450
151 lIleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
451 ATTGTTATATTTGACAGCAAGTGAATTATAGGTTATTCATCATCGATGA 500
167 pAsnPheIleLysLysAsnThrProGlnGlyPheLysLysSerPheI 184
501 CAACCTTATTTAAAAAACACCTCCACAAAGTTTCAAGAGATCAAGTTT 550
184 yThrValGluValAsnArgLysGluGluValAlaLeuIleMetAsnSer 200
551 AAACGTGAGAAAGTTAACCGCAAGAACAAAGTTGCTTATATATCAATCT 600
201 SerGlySerThrGlyLeuProLysGlyValGluLeuThrHisGluAsn 217
601 TCGGCTTCAACCGGTTTCCAAAAAGTTGTCACTTACTCATATAAATTT 650
217 uValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGluVal 234
651 GGTGATCGATTTTCTCAGCTGAGATCAATTTATGAAACCAAGTTT 700
234 eProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
701 CACGAGCAGCGCATTTTAACTGATACCATTCATCATGATGTTTGTG 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetIle 267

```

```

|||||
751 ATGTTACTACTTGTAGGCTATCTAGCTGTGTTGCTATTCATCTT 800
267 urhrlyspheaspjlugluhrPheleuylsThrleuGlnsPtyrLysC 284
801 AACGAATTGACGAGAGACTTTTAAACACTGCAGATTCAAAAT 850
284 ysserSerValIleLeuValProthrLeuphealIleLeuasnrgSer 300
851 GTTCAGCGGTATTTCTTACGACACTTCTTGCAATCTTATATGAAAGT 900
301 GluleuLeuaspLysTyrAspleuSerAsnLeuValGluIlealSerGI 317
901 GAATTACTCGATAAATATGATTTATCAAAATTAGTTGAATTCGATCTCG 950
317 yGlyAlaProleuSerLysGluIleGlyGluAlaValAlaArgPheA 334
951 CGGAGCACCTTTATCTTAAGAAATGCTGAACTGCTGCTAGACGTTTA 1000
334 snleupProGlyValArgGluGlyTyrGlyLeuThrGluThrThrSerAla 350
1001 ATTACCGCGGTCTGTCAGAGGCTATGCTTTAACAGAAACACCTCTGCA 1050
351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
1051 ATTATATACACCGGAGCGGATGATAAACAGTGCTCTGCAAAAGT 1100
367 lValProleuPheLysAlaLysValIleaspLeuAspThrLysLysThrL 384
1101 TGTGCCATTATTTAAAGCAAAAGTTATCGATCTTGATCTATAAAACTT 1150
384 euGIYProAsnArgArgGlyLysValLysValLysGlyPrometLeuMet 400
1151 TGGGCCCGACAGAGCTGAGAAAGTTGTGTAAAGGGCTCTATGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleaspGluGI 417
1201 AAAGGTTATGATATATCCAGAACCAACAGAAATCATATGATGAAGA 1250
417 uGIYTrpLeuHisThrGlyAspIleGlyTyrTyrAspLysGluLysHisP 434
1251 AGGTGTGTTGCACACGAGATATTGGGATATACGATGACAGAAAAACATT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
1301 TCTTATCGTGATCGTTGAAAGTCTTTATCAAAATACAAAGATATCA 1350
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
1351 GTACACACTCTGATTAAGATCTGTTTGCACACATCCAAATATTTT 1400
467 eaSPAlaGlyValAlaGlyValProaspProIleAlaGlyLysLeuProG 484
1401 TGATGCCGCGGTGCTGCGTCCAGATCTTATAGCTGCTGACCTCCGG 1450
484 lYAlaValIleValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
1451 GAGCTGTTGTTGTTACTTAAGAAAGAAATCTATGACTGAAAAAGAAAGA 1500
501 MetaspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgLysGI 517
1501 ATGATTCAGTTCAGTCAAGTTCAAAATGCAAAACGTTTGGCTGTGCTG 1550
517 yValAlaPheValAspGluValProLysGlyLeuThrGlyLysLysSpG 534
1551 TGTCCGTTTTTGTGACGAGACTAAAGCTCCTACGTGTAATAATTGACG 1600
534 lYlYsAlaIleArgGluLeuLeuLysLysProValAlaLysMet 548
1601 GTAAAGCAATTAGAGAAATACTGAAGAAACAGTTGCTAAGATG 1644

```

seq_name: /SIDS2/gcdata/geneseq/geneseqn/NA1993.DAT.AA034745

seq_documentation_block:

```

ID   AA034745 standard; cDNA to mRNA; 1644 BP.
XX
AC   AA034745;
XX
DT   24-MAY-1993 (first entry)
XX
DE   Thermostable luciferase (wild-type).
XX
KW   Firefly; Luciola cruciata; GEMUT; Luciola lateralis; HEIKE;
XX   mutant; assay; ss.
XX
OS   Luciola lateralis.
XX
PN   EP524448-A.
XX
PD   27-JAN-1993.
XX
PF   26-JUN-1992; 92EP-0110808.
XX
PR   27-JUN-1991; 91JP-0157117.
XX   29-NOV-1991; 91JP-0317064.
XX   22-NOV-1992; 92JP-0131057.
XX
PA   (KIKR ) KIKROMAN CORP.
XX
PI   Eliehl N, Naoki K;
XX
DR   WPI: 1993-028553/04.
XX   P-PSDB: AAR30803.
XX
PT   New DNA sequence of a thermostable luciferase - has the aminoacid
XX   at position-217 of Luciola cruciata or lateralis luciferase
XX   replaced by a hydrophobic aminoacid
XX
PS   Claim 1-3; Page 29-30 + 21-24; 33pp; English.
XX
CC   A DNA sequence of a thermostable luciferase of a firefly, which
XX   encodes the amino acid sequence of a wild-type firefly luciferase in
XX   which an amino acid at the 217 position or an amino acid at the
XX   position equiv. to the 217 position of the luciferase of Luciola
XX   cruciata (GEMUT firefly) of Luciola lateralis (HEIKE firefly) is
XX   replaced by a hydrophobic amino acids, esp. Ile, Leu or Val, is
XX   claimed. The mutant luciferase is identical in properties to the
XX   wild-type luciferase except that it is stable when heated to high
XX   temps., e.g. 50 degrees C. The luciferase can be used in assays for
XX   e.g. ATP.
XX
SO   Sequence 1644 BP; 529 A; 262 C; 349 G; 504 T; 0 other;

alignment_scores:
      Quality: 2814.00      Length: 548
      Ratio: 5.144          Gaps: 0
      Percent Similarity: 99.818      Percent Identity: 99.635

alignment_block:
US-09-581-241-4 x AA034745 ..
Align seg 1/1 to: AA034745 from: 1 to: 1644

1 MetGlnsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17
|||||
1 ATGGAAGAACATGAGAGATGAAATATTTGTTGTTGCTGTAACCATTT 50
17 eTyrProIleGluGluGlySerAlaGlyAlaGluLeuArgLysTyrMetA 34
|||||
51 TTACCTATTGAGAGAGATGCTGCGACACAATTGCCGCAAGTATATG 100
34 sPArgTyrAlaLysLeuGlyAlaIlealPheThrAsnAlaLeuThrGly 50
|||||
101 ATCGATATGCAAAACTTGAGCAATTCGCTTTACTAAGCACTTACCGGT 150

```

51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysLeuGly 67
 151 GTCGATTTATAGCTACGCCGAACTACTTAGAAAAATCATGCTGTCTAGAGA 200
 67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
 201 GGCCTTAAGAAATTAATGTTGGTTGGTTGATGAGAAATTCGCTATGCA 250
 84 eRGluAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
 251 GTGAAACCTGGAAGAATCTTATTCCTGTATACCGGTTATATTATA 300
 101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
 301 GGTGTCGGTGGCTCCAACTAATGAGATTACACCTACGTAAGTAATGGT 350
 117 LHisSerLeuGlyIleSerLysProThrIleValIlePheSerSerLysG 134
 351 TCACAGTTTAGCATCTCTAAGCCACACATTTGATTTACTTTAAAG 400
 134 LysLeuAspLysValIleThrValGluLysThrValIleAlaIleLysThr 150
 401 GATTAACTAAAGTTTAACTGACAAAAACGTAACCTGATTAATAAC 450
 151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAs 167
 451 ATGTGATATATGGACAGCAAGTGGATTATAGAGTTATCAATCCATGA 500
 167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerPheL 184
 501 CAACCTTTATTAATAAAACACCTCCAAAGGTTTCMAAGATCAAGATT 550
 184 ystThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
 551 AAACGTGTAAGTAAACGCAAGACAGACTGCTTTATATATGACTCT 600
 201 SerGlySerThrGlyLeuProLysGlyValGluLeuThrHisGluAsnLe 217
 601 TCGGTTTAAACGGGTTTGCAAAAGGTTGCACTACTCATCATGAATATGC 650
 217 uValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGluValS 234
 651 AGTCACTGATTTCTCTACGCTACAGATCCATTTATGCAACCAAGTTT 700
 234 eRProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
 701 CACCAAGGACGGCTATTTAACTAGTACTACCATTCATCATGTTTGGT 750
 251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleAlaMetIle 267
 751 ATGTTTACTACTTAGGCTATCTTAACCTGTGTTTGTGATTTGTCATGTT 800
 267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
 801 AACGAAATTTGACGAAGACACTTTTAAAAACACGCAAGATTCCAAAT 850
 284 yssSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
 851 GTTCAAGCGTTATCTGTATACGACTTTTGTTCGAAATCTTAATAGAGT 900
 301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
 901 GAATTACTCGATAAATATGATTTATCAAAATTAAGTTGAATTCGATCTGG 950
 317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheA 334
 951 CGAGAGCACTTATCTAAAGAAATGTTGAGAGCTGTTCGTAGAGCTTTTA 1000
 334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
 1001 ATTATACCGGTTGTCGCAAGGCTATGTTTAAACGAAACACTCTGCA 1050
 351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367

1051 ATTTATATACACCGGAGGAGGATTAACCAAGGCTCTTGGCAAGT 1100
 367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysThrL 384
 1101 TGTCCCATTTATTTAAACCAAAAGTTATTCATCTGTGACTAAAAAATTT 1150
 384 eugLysProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
 1151 TGGGCCCGAACAAGAGTGGAGAAATTTGTTAAAGGTCCTATGCTATG 1200
 401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGlu 417
 1201 AAAGGTTATGTAGATTAATCAGAACCAACAGAAATCATATGATGAAGA 1250
 417 uGlyTyrPheHisThrGlyAspIleGlyTyrTyrAspGluLysHisP 434
 1251 AGGTTGGTTGCACACAGAGATTTGGGTATTCGATGAGAAAAACATT 1300
 434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
 1301 TCTTTATCGTGATGCTTTGAAGCTTTATCAATATCAAAAGATATCA 1350
 451 ValProProAlaGluLeuGlnSerValLeuLeuGlnHisProAsnIleP 467
 1351 GTACCACTGCTGAATTAAGAAATCTGTTCTTTGCAACATCCAAATATTT 1400
 467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuPro 484
 1401 TCATCCCGCGCTGCTGGCTGCAGATCCATATGCTGCTGAGCTCCG 1450
 484 lValAlaValAlaLeuLysLysGlyLysSerMetThrGlyLysGluVal 500
 1451 GAGCTGTTGTTGTAAGTGAAGAAATCTATGACTGAAGAAAGAA 1500
 501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
 1501 ATGATTACGTTGCTAGTCAAGTTTCAAAATGCAAAAGCTTTGGCTGG 1550
 517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
 1551 TGTCCGTTTGTGGACGAAGTACTTAAAGGCTCTCACTGTTAAATATGAC 1600
 534 lLysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
 1601 GTAAAGCAATTAGAAGAAATCAACAAACAGTTGCTAGATG 1644
 seq_name: /sids2/gcgdata/geneseq/NA1997.DAT: AAT63269
 seq_documentation_block:
 ID AAT63269 standard; DNA: 1908 BP.
 AC AAT63269;
 XX
 DT 20-MAY-1997 (first entry)
 XX
 DE Gene for firefly luciferase conjugated with 87 aa peptide at 3' end.
 XX
 KM Fusion protein; firefly; luciola lateralis; luciferase; biotinylation;
 KM wild type; E.coli; bioluminescence assay; quantification; ligand;
 KM receptor; ds.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT 1..1641
 FT CDS
 FT /*tag= a
 FT /note= "encodes firefly portion of fusion protein"
 FT 1642..1647
 FT /*tag= b
 FT /note= "sequence encoded by linker used to insert
 FT sequence encoding biotin peptide to generate
 FT fusion gene"

```

FT      CDS      1648..1908
FT      /**tag=
FT      /note= "encodes biotinylated peptide portion of fusion
FT      protein"
PN      JP08308578-A.
XX      26-NOV-1996.
XX      24-APR-1995; 95JP-0039857.
XX      14-MAR-1995; 95JP-0054625.
XX      27-JUL-1994; 94JP-0193798.
XX      (KIKK ) KIKKOMAN CORP.
XX      WPI; 1997-059697/06.
XX      P-PSDB; AAM12395.
DR      Fusion protein comprising firefly luciferase and biotinylated
DR      peptide - useful in a bio-luminescent analytical method for
PT      quantifying ligands
PT      Example 7; Page 11-12; 13pp; Japanese.
XX      This is the nucleotide sequence encoding a novel fusion protein which
XX      comprises the firefly (luciola lateralis) luciferase protein and an 87
XX      amino acid biotinylated peptide designated the E.coli biotin carboxy
XX      carrier protein (BCCP-87). The fusion gene was generated by firstly
XX      amplifying the sequence encoding the E.coli (BCCP-87) from the E.coli
XX      genome and inserting the resultant sequence into the 3' end of the
XX      firefly luciferase gene in plasmid pHLf230. The biotin peptide sequence
XX      replaces the C-terminal Met residue of the wild type luciferase sequence.
XX      The plasmid was then introduced into E.coli JM101 for production of the
XX      fusion protein. The novel protein can be used in bioluminescence assays
XX      to quantify luciferase ligands which may modulate binding of luciferase
XX      to its receptor.
XX      Sequence 1908 BP; 598 A; 329 C; 426 G; 555 T; 0 other:
XX
alignment_scores:
Quality: 2814.00      Length: 547
Ratio: 5.144          Gaps: 0
Percent Similarity: 100.000      Percent Identity: 99.817
alignment_block:
US-09-581-241-4 x AAT63269 ..
Align seg 1/1 to: AAT63269 from: 1 to: 1908
1 MetGluAsnMetGluAsnAspGluAsnIleValTYrGlyProGluProPh 17
1 ATGGAAGAAACATGAGAACGATGAAATATTTGTATGGTCTGAAACCATTT 50
17 eTYrProIleGluGluSerAlaGlyAlaGluIleuArGlySTYrMetA 34
51 TTACCTATTGGAAGAGGATCTGCTGAGCACATTTGCCAAGTATATGG 100
34 sPArGTYrAlaIysLeuGlyAlaIleAlaPheThrAsnAlaIeuThrGly 50
101 ATCGATATGCAAAACTTGAGCAATTCCTTTACTAACCGACCTTACCGGT 150
51 ValAspTYrThrTYrAlaGluTYrLeuGluIlySerCysCysLeuGly 67
151 GTGCAATTATACGTCGCCGAAATCTTGAAGAAATCATGCTGCTAGAGACA 200
67 uAlaLeuIyAsnTYrGlyLeuValValAspGlyArGlyIleAlaLeuCys 84
201 GGCCTTAAAGATTAATGCTTGGTGTGTGATGGAAGAAATTCGCTATGCA 250
84 eRgiuAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100

```

```

251 GTGAAGAACTGGAAGAGTCTTTATTCCTGTATTAAGCGGTTATTATA 300
101 GlyValGlyValAlaProThrAsnGluIleTYrThrIleuArGlyIleuVal 117
301 GGTGCGGTGTGGCTCCAACTAATGAGATTTTACACTTACTGTAATGGT 350
117 lHisSerIleuGlyIleSerIlySProThrIleValPheSerSerIlySG 134
351 TCACAGTTTAGCATCTCTTAAGCCACACATTTGTATTGATGTTCAAAAAAG 400
134 lYLeuAspIlyValIleThrValGluIlyThrValThrAlaIleIlyThr 150
401 GATTAGATAAAGTTATTAAGTACAAAAACGTAACGCTATTATTAACCC 450
151 lIleValIleIleuAspSerIlyValAspTYrArGlyTYrGlnSerMetAs 167
451 ATTGTTATATTGGACAGCAAGTAGATTAAGAGTTATTCATATCATGGA 500
167 pAsnPheIleIlyLysAsnThrProGlnGlyPheIlySGIlySerSerPheI 184
501 CAACCTTATTAAAAAACAACCTCCACAAGGTTTCAAGGATCAAGTTTGA 550
184 ySThrValGluValAsnArGlySGluGluValAlaLeuIleMetAsnSer 200
551 AAACCTGTAGAACTTAACCGCAAAAGCAAGTGTCTTATTAATGAACCT 600
201 SerGlySerThrGlyLeuProIlySGIlyValGluIleuThrHisGluAsn 217
601 TCGGGTTCAACCGGTTGGCCAAAGAGTGCAACTTACTCATGAAATTT 650
217 uValThrArGpHeSerHisAlaArGAsProIleTYrGlyAsnGluValS 234
651 GGTACAGCGGTTTTTTCACGCTAGAGATCCAAATTTGGAACCAAGTT 700
234 eRProGlyThrAlaIleuThrValValProPheHisIlySGIlyPheGly 250
701 CACCAAGCAGCGCTATTTTACTGTAGTACCATTCATCATGTTTGGT 750
251 MetPheThrThrIleuGlyTYrIleuThrCysGlyPheArGlyIleValMet 267
751 ATGTTTACTACTTAGGCTATTAACCTTGGTTCGTTCGTATTCATGTT 800
267 uThrIlySPheAspGluGluThrPheIleuIlySThrIleGluAspTYrLysC 284
801 AACCAAAATTTGACGAAAGACCTTTTAAACAACGTCGCAAGATTACAA 850
284 ySSerSerValIleLeuValProThrIleuPheAlaIleuAsnArGSer 300
851 GTTCAGCGTATTCTTGTACCGACCTTGTTCGAATTCCTTAATAGAA 900
301 GluLeuLeuAspIlyTYrAspIleuSerAsnLeuValGluIleAlaSerG 317
901 GAATTACTCGATTAATATGATTTATCAAAATTTAGTTAAATTTGATCTG 950
317 yGIValAlaProIleuSerIlyIleGlyGluAlaValAlaArGpPheA 334
951 CGAGACACCTTATCTTAAGAAATTTGGTGAAGCTGTTGCTAGACGTTT 1000
334 snLeuProGlyValArGlnGlyTYrGlyLeuThrGluIlyThrSerAla 350
1001 ATTTACCGGTTGTCGTCAAGGCTATGTTTAAAGAAACAACCTGTGCA 1050
351 lIleIleIleThrProGluIlyAspAspIlySProGlyAlaSerGlyIlyS 367
1051 ATTATTATTCACCGCGAAGGCGATGATTAACACAGGCTGTTGCGCAAG 1100
367 lValProIleuPheIlyAlaIlyValIleAspIleuAspThrIlySThrL 384
1101 TGTGCAATTATTAAAGCAAAAGTTATGCACTTATCTATTAATAAACTT 1150
384 euGIlyProAsnArGArGlyGluValIlyValIlySGIlyProMetIleu 400
1151 TGGGCCCGAAGACGAGTGAGAAAGTTTGTGTAAGGCTCTATGCTTATG 1200

```

```

401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleLeuSgLucI 417
    |||
1201 AAGGTATAGTATATATCCACAGCAACAGAAATCATATGATGAGA 1250
    |||
417 uGlyTTrpleuHsthThGlyAspIleGlyTyrTyrAspGluGlyHisP 434
    |||
1251 AGGTGGTTGGACACAGAGATATGGTATATACGATGAGAAACATTT 1300
    |||
434 hepHeilleValAspArgLeuLysSerLeuIleLeuTyrLysGlyTyrGln 450
    |||
1301 TCTTATAGTGGATGCTTGAAGCTTTAAATCAATACAAAGGATATCAA 1350
    |||
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
    |||
1351 GTACCACTGCTGAATATGAAATCTGTTCTTTCGAACATCCAAATATTTT 1400
    |||
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
    |||
1401 TGATGCCCGCGCTTGTGCTGCCATCCAGATCTATAGCTGGTACGCTCCG 1450
    |||
484 lYAlaValAlaValLeuLysGlyLysSerMetThrGluLysGluVal 1500
    |||
1451 GAGCTGTGTGTACTGAAAGAAAGAAATCTATGACTGAAAAAGAGTA 1500
    |||
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyG 1517
    |||
1501 ATGATTTACGTTGCTAGTCAAGTTCAAAATGCAAAACGTTGGCTGATG 1550
    |||
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
    |||
1551 TGTCGGTTTGTGGACGACGACTTAAGCTCAGCTGCTGTAATAATTGACG 1600
    |||
534 lYlYsAlaIleArgGluIleLeuLysLysProValAlaLys 547
    |||
1601 GTAAAGCAATTTAGAAATATCTAGAGAAACCAAGTTGCTAAG 1641
    |||
seq_name: /SID52/gc9data/geneseq/geneseqn/NA1998.DAT:AAV32467
seq_documentation_block:
ID AAV32467 standard; DNA; 1920 BP.
AC AAV32467;
XX
XX 10-SEP-1998 (first entry)
XX
XX Luciferase-lysostaphin fused gene.
XX
XX Luciferase-lysostaphin fused gene; chimeric; bioluminescence; ds.
XX
XX Chimeric - Staphylococcus simulans.
XX
XX Chimeric - Luciola lateralis.
XX
XX Key Location/Qualifiers
XX CDS 1..1920
XX /*tag=
XX /product= "Luciferase-lysostaphin fusion protein"
XX /note= "Nucleotides 1-1647 encode luciferase
XX protein while nucleotides 1648-1920 encode
XX lysostaphin; CDS does not contain stop codon"
XX
XX JPI0150991-A.
XX
XX 09-JUN-1998.
XX
XX 25-NOV-1996; 96JP-0328042.
XX
XX 25-NOV-1996; 96JP-0328042.
XX
XX (KIKK ) KIKKOMAN CORP.
XX
XX WPI; 1998-379994/33.
XX
XX P-PSDB; AAW4856.
XX

```

```

XX
XX New Luciferase-lysostaphin fused protein - useful for
XX bioluminescence analysis
XX
XX PS Disclosure; Pages 6-7; 10pp; Japanese.
XX
XX CC The invention claims for a luciferase-lysostaphin fused gene which
XX encodes a fused protein in which a peptide part consisting of
XX luciferase, from Luciola lateralis, is connected to a peptide part
XX consisting of lysostaphin from Staphylococcus simulans. The method
XX described by the invention can be used to prepare a luciferase-
XX lysostaphin fused protein efficiently. The fused protein is useful
XX for bioluminescence analysis.
XX
XX SQ Sequence 1920 BP; 627 A; 310 C; 407 G; 576 T; 0 other;

alignment_scores:
    Quality: 2814.00      Length: 547
    Ratio: 5.144          Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 99.817

alignment_block:
US-09-581-241-4 x AAV32467
..
Align seg 1/1 to: AAV32467 from: 1 to: 1920

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17
    |||
1 ATGGAACATATGAGACAGATGAAATATGTGATATGCTGTAACCATTT 50
    |||
17 eTyrProIleGluGluGlySerAlaGlyAlaGluLeuArgLysTyrMet 34
    |||
51 TTACCTATTGACAGGAGATCTGCTGGAGCAATGTGGCAAGTATATG 100
    |||
34 sPARGTyrAlaLysLeuGluValAlaIleAlaPheThrAsnAlaLeuThrGly 50
    |||
101 ATCGATATGCAAAACTTGGACCAATGCTTTACTTAACGCCATTAACCGGT 150
    |||
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysLeuGlyG 67
    |||
151 GTCGATTATACGTAGCGCGCAATCTTACGAAACATCAATGCTGTACGAGA 200
    |||
67 uAlaLeuLysAsnTyrGlyLeuValAlaAspGlyArgIleAlaLeuGly 84
    |||
201 GGCCTTAAAGAAATATATGTTGGTTGTTGATGAGAAATTCGTTATGCA 250
    |||
84 eArgLysCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
    |||
251 GTGAACCTGTGAAGAGTCTTATCTCTGTAATGACCGGTTATTATATA 300
    |||
101 GlYAlaGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
    |||
301 GGTGCGGTGGTGGCTCCCACTAATGAGATTTACACTCTACGTAATTTGCT 350
    |||
117 lHisSerLeuGlyIleSerLysProThrIleValIleSerSerLysG 134
    |||
351 TCACAGTTTAGCATCTCTAAGCCCAACATTTGATTTAGTTCAAAAAG 400
    |||
134 lYLeuAspLysValIleThrValGluLysThrValThrAlaIleLysThr 150
    |||
401 GATTAGATAAGTTATTAAGTACCAAAAAACGGTAATCTCTATTAAAC 450
    |||
151 lIleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
    |||
451 ATGTATATATTTGACAGCAAGTGATTTAGAGTTATCAATTCATGGA 500
    |||
167 pAsnHeilleLysLysAsnThrProGlnGlyPheLysGlySerSerPhe 184
    |||
501 CAAGTTTATTAATAAAACATCCACAAAGTTCAAAAGATCAAGTTTAA 550
    |||
184 ystThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
    |||

```

```

551 AAACGTAGAGTTAACGCCAAGAACAACTGCTTATTAATGAACCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLe 217
|||||
601 TCGGGTTCAACCGGTTGGCAAAAGGTGTGCAACTACTCATGAAATATT 650
217 uValThrArgPheSerHisAlaArgAspIleTyrGlyAsnGlnValS 234
|||||
651 GGTACAGCGCTTTTTCACGCTAGAGATCCAAATTATGGAACCAAGTTT 700
234 expProGlyThrAlaIleLeuThrValValProPheHisGlyPheGly 250
|||||
701 CACCAGGACGCGTATTATTTACTGTAGTACATTCATCATGGTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
|||||
751 ATGTTACTACTTACTTAGCTATCACTGTGCTTTCGTTATGTGCATGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
|||||
801 AACGAAATTTGACGAAGACGCTTTTAAACACACTGCAGATTACAAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
|||||
851 GTTCAACCGTTATCTTGTACGACCTTGTTCCAATTCCTTATAGACGT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGlnIleAlaSerG 317
|||||
901 GAATTACTCGCATTAATGTGATTTATCAAAATTTAGTGAATTCATCTCG 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
|||||
951 CGAGACACCTTTATCTTAAGAAATGTGTGAAGCTGTTCGTAAGCGTTT 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrSerAla 350
|||||
1001 ATTACCGGGGTTCGTCAGGCTATGTTTAAACAGAAACACCTTGCA 1050
351 IleIleIleThrProGluGlyAspAspLysProGlyValSerGlyLysVa 367
|||||
1051 ATTATTTATCACCGGAGGCGAGATGAACCAAGGTCTTCTTGCAAGGT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThr 384
|||||
1101 TGTGCCATTTATTAAAGCAAAAGTTATCGATCTGTATCAAAAAAATT 1150
384 euGlyProAsnArgArgGlyValCysValLysGlyPrometLeuMet 400
|||||
1151 TGGGCCCGAAGACGAGCGAGAAAGTTGTGTAAGGGTCTATGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgLutIleIleAspGluG 417
|||||
1201 AAAAGTTATGTAGATTAATCCAGAAACCAAGAAATCATAGATGAAGA 1250
417 uGlyThrLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
|||||
1251 AGGTTGGTTGCACACAGAGATATTGGGTATTACGATGAAGAAAAAATT 1300
434 hepHeIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
|||||
1301 TCTTTATCGTGGATCGTTTGAAGTCTTTAATCAATACAAAGGATATCAA 1350
451 ValProProAlaGluLeuGluLysSerValLeuLeuGlnHisProAsnIlePh 467
|||||
1351 GTACACCTGCTGATTAATGATCTGTTCTTTTGCACATCAAAATATTATT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyLutLeuProG 484
|||||
1401 TGATCCGCGCGTTGCTGCGCTTCGATCCATACCTGCTGAGCTTCGCG 1450
484 lValAlaValValValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
|||||
1451 GAGCTGTGTTGTACTTGAAGAAAGAAATATCTATACAGAAAAAAGAACTA 1500

```

```

501 MetAspTyrValAlaIleSerGlnValSerAsnAlaLysArgLeuArgLysI 517
|||||
1501 ATGGATTACGTTGCTAGTCAAGTTTCAATCAAAACGTTTGGCGGTGG 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
|||||
1551 TGTCCGTTTGTGGAGCAAGTACCTAAAGTCTCAGTGTAAATTTGACG 1600
534 lLysAlaIleArgGluIleLeuLysLysProValAlaLys 547
|||||
1601 GTAAAGCAATTAGAGAAATTAAGTGAAGAAACCAAGTTGCTAG 1641
seq_name: /SIDIS2/gcgdata/geneseq/geneseqn/NA1990.DAT:AA003257
seq_documentation_block:
ID AA003257 standard; DNA: 1644 BP.
XX
AC AA003257:
XX
DT 22-JUL-1990 (first entry)
XX
DE Luciola lateralis recombinant luciferase gene.
XX
KW Luciferase; enzyme; ATP; plasmid pHLf7; firefly; ss.
XX
OS Luciola lateralis.
XX
FH Key Location/Qualifiers
FT CDS 1..1644
FT /*tag= a
FT /product=luciferase
FT /EC_number=EC-1.13.12.7
XX
PN EP353464-A.
XX
PD 07-FEB-1990.
XX
PF 30-JUN-1989; 89EP-0111958.
XX
PR 22-DEC-1988; 88JP-0162402.
PR 01-JUL-1988; 88JP-0162402.
XX
PA (KIKK) KIKKOMAN CORP.
XX
PI Tatsumi H, Kajiyama N, Nakano E;
XX
DR WPI: 1990-038240/06.
XX
DR P-PSDB; AAR03731.
XX
PT New gene encoding luciferase -
PT derived from Luciola lateralis and expressed in and purified
PT from Escherichia coli genus cells.
XX
PS Claim 3; page 24; 43pp; English.
XX
CC This enzyme is useful for determining ATP levels. It is
CC contained in plasmid pHLf7.
XX
SQ Sequence 1644 BP; 529 A; 264 C; 348 G; 503 T; 0 other;

```

```

alignment_scores:
  Quality: 2811.00      Length: 548
  Ratio: 5.139          Gaps: 0
  Percent Similarity: 99.818  Percent Identity: 99.453

```

alignment_block:

US-09-581-241-4 x AA003257 ..

Align seq 1/1 to: AA003257 from: 1 to: 1644

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17

```

|||||
1 ATGGAAAAATGAGACGATGAAAAATATGCTATGCTGACCACTT 50
17 eYrProIleGluGluGlySerIleGluValAlaIleuArgLysTyrMet 34
51 TTACCTATTGAGAGGATCTGCTGAGACCAATTCGCAAGATATATGG 100
34 spArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 ATCGATACGCAAACTTGAGCAATTCCTTTTACTAACGCACTTACCGGT 150
51 ValAspTyrThrTyrAlaGluTyrIleuGluLysSerCysLeuGlyGly 67
131 GTCGATTTACGTACGCGCAATCTTGAAGAAAATCATGCTGTACAGAGA 200
67 uAlaLeuLysAsnTyrGlyLeuValAlaSpGlyArgIleAlaLeuCys 84
201 GGCCTTAAGAAATTTATGTTGGTTGTGATGGAAGAAATTCGTTATGCA 250
84 eArgLysAsnCysGluGluPhePheIleProValIleuAlaGlyLeuPheIle 100
251 GTGAAAACTGTGAAGAAATTTCTTATTCCTGTATTAGCCGGTTATTATATA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
301 GGTGCGGTGCTGCTCCACTAATGAGATTACACTGACGTGAATTTGTT 350
117 LHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysGly 134
351 TCACAGTTAGGCATCTCTAAGCCAAATGTATTGATTCTTAAGAAAG 400
134 LysLeuAspLysValIleThrValGlnLysThrValIleThrAlaIleLysThr 150
401 GATTAGATTAAGTTATTAAGTGTACAAAAAAGGTAAGTCTGATTAAAAACC 450
151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
451 ATGTATTATTTGACAGCAAAAGTGATTAAGAGTTATCAATCCATGCA 500
167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheIle 184
501 CAACCTTTATTAAAAAACAACCTCCACAAGGTTTCAAAAGGATCAAGTTTGA 550
184 yThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
551 AAATGTAGAAGTTAACCGCAAGAACAAAGTTGCTTATTATATGAACTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLe 217
601 TCGGGTTCAACCGGTTGCCAAAAAGGTGCAACTTACATCATGAAAATGC 650
217 uValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnVal 234
651 AGTCACATAGATTTCTCAGCTAGAGATCCAAATTTATGGAACCAAGTTT 700
234 eArgProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
701 CACCAGGACGCGCTATTATTAACTGATGACATTCATCATGTTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgGlyLeuMetLe 267
751 ATGTTTACTACTTATAGCTATTAACCTTGTTGTTCTGATTGTCTATGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAsnTyrLysC 284
801 AACGAATTTGAGAAAGACATTTTAAAAAACACTGCAGATTTACAAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTTCAAGCGCTATTCTGTACGACTTTGTTGCATTTCTTAATAGAACT 900
301 GluLeuLeuAspLysTyrThrLeuSerAsnLeuValGluIleAlaSerGly 317
|||||
901 GAATTACTCGATAATATGATTTATCAAAATTTAGTTGAAATTCGATCTGG 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
951 CGGAGCACCTTTATCTPAAAGAAATGTGTAAGCTGTGTGACGCTTTTA 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrSerAla 350
1001 ATTTACCGGTGTGTCGAAGGCTATGTTTAAACAGAAAACCTCTGCA 1050
351 IleIleThrProGluGluLysAspLysProGlyAlaSerGlyLysVal 367
1051 ATTATTATCACACCGGAAGCGATGATTAACCAAGGTGCTTCTGCAAACT 1100
367 ValProLeuPheLysAlaLysValIleAspLeuAspThrLysThrL 384
1101 TGTGCATTAATTTAAAGCAAAAGTTATCATGCTGTGATCTAAAAAAA 1150
384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
1151 TGGGCCCGAAGACAGCTGAGAAATTTGTGTAAGGCTCTATGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGlu 417
1201 AAAGTTATGTAGATTAATCCAGAACAGAACAGAAATCATGATGAGGA 1250
417 uGlyTyrPheHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
1251 AGTTTGTGTCACACGAGAGATTTGGGATTTACATGAGAAAGAAACAT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
1301 TCTTTATCGTGATCGCTTGAAGTCTTTATCAATTAACAAAGATATCA 1350
451 ValProProAlaGluLeuGlnSerValLeuLeuGlnHisProAsnIlePh 467
1351 GTACCACTGCTGAATTAAGATCTGTTTGGCAACATCCAAATATATTT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
1401 TGATGCGGCGCTGCTGCGGCTCCAGATCCTTAAGTCTGACGCTTCGG 1450
484 ValAlaValValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
1451 GAGCTGTGTTGTTACTGAAAAAGGAAATCATGATGAAAAAGAACTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyGly 517
1501 ATGGAATTACGTTGCTAGTCAAGATTCAAAATGCAAAACGTTGCGGTGG 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAsp 534
1551 TGTCCGTTTTGGGAGCAACTAAGGTCTCAGCTGTGTAATTTGACG 1600
534 LysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
1601 GTAAAGCAATTAAGAAATTAAGTGAAGAAACCAAGTTGCTTAAGATG 1644

```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:AA63268

seq_documentation_block:

ID AA63268 standard; DNA: 1704 bp.

AA63268:

20-MAY-1997 (first entry)

Firefly Luciferase conjugated with 23 aa peptide at 5' end.

Fusion protein: firefly: lucifera lateralis: luciferase: biotinylation;
 wild type: E.coli: bioluminescence assay: quantification; ligand;
 receptor; ds.

XX


```

203 SerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLeuValTh 219
|||||
667 TCACCGCGTTGGCCAAAGGTGGCAACTTACTCATGAAATTTGGTGCAC 716
|||||
219 FatpHeserHisAlaArgAspProIleTyrGlySngInValSerProg 236
|||||
717 GCGTTTTTCTCAGCCTAGAGATCCAAATTTATGMAAACCAAGTTTCACCG 766
|||||
236 LThrAlaIleLeuThrValValProPheHisHisGlyPheGlyMetPhe 252
|||||
767 GCACGCGCTATTTAACTAGTACATCCATCCATGCTTTGGTATGTTT 816
|||||
253 ThrThrLeuGlyTyrIleuThrCysGlyPheArgIleValMetLeuThrly 269
|||||
817 ACTACTTGGCTATCTACTGTGGTTTCCGATTTGTCATGTTAAACGAA 866
|||||
269 sPheAspGluGluThrPheLeuLysThrLeuGluAspTyrLysCysSers 286
|||||
867 ATTTGACGAAGAGACTTTTAAAAACACTGCAGATTTACAAATGTTCAA 916
|||||
286 erValIleLeuValProThrLeuPheAlaIleLeuAsnArgSerGluLeu 302
|||||
917 GCGTATTTCTTGTACCGACTTTGTTGCAATCTTAATAGAGTGAATTA 966
|||||
303 LeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGlyGlyAl 319
|||||
967 CTGCATTAATATGATTTATCAAAATTTAGTTGAAATTCATCTGGCCGAGC 1016
|||||
319 aProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheAsnLeuP 336
|||||
1017 ACCTTATCTAAAGAAATGTTGGTAGAGCTGTGTACAGCTTTTAATTTAC 1066
|||||
336 roGlyValAlaArgGlnGlyTyrGlyLeuThrGluThrSerAlaIleIle 352
|||||
1067 CGGGTGTCTTCACAGGCTATGTTTAAACAGAAACACCTCTGCAATTAAT 1116
|||||
353 IleThrProGluGlyAspAspLysProGlyAlaSerGlyLysValValPr 369
|||||
1117 ATCACACCGGAGAGCGCATGTATTAACCGAGTGTCTTGGCAAAATGTGGCC 1166
|||||
369 OleuPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGlyP 386
|||||
1167 ATTTATTTAAAGCAAAAGTATTCGATCTTGATACATAAAAAAATTTGGGCC 1216
|||||
386 roAsnAlaGlyGlyGlyValCysValLysGlyProMetLeuMetLysGly 402
|||||
1217 CGAACACAGCGGAGAGTTGTGTAAAGGTCCTATGCTTATGCAAGGT 1266
|||||
403 TyrValAspAsnProGluAlaThrArgGluIleIleAspGluGluGlyTr 419
|||||
1267 TATGTAGTATTAATCCAGAACACAGCAAGAAATCTATAGTAAAGAGTTG 1316
|||||
419 PLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisPhePheI 436
|||||
1317 GTTCACACAGAGAGATATGTGGTATTCAGATGAAGAAAAAATTTCTTTTA 1366
|||||
436 leValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGlnValPro 452
|||||
1367 TCGTGAGATCGTTGAAGCTTTTAATCAAAATPACAAAGATATCAAGTACCA 1416
|||||
453 ProAlaGluLeuGluSerValIleuLeuGlnHisProAsnIlePheAspAl 469
|||||
1417 CCTGTGATTAAGATCTGTTCTTTGCAACATCAATATTTTGTATGTC 1466
|||||
469 agIyValAlaGlyValProAspProIleAlaGlyGluLeuProGlyLay 486
|||||
1467 CGGGCTTGTCTGCGTCCAGATCTATAGCTGTGAGCTTCCGGAGACTG 1516
|||||
486 alValValLeuLysGlyLysSerMetThrGluLysGluValMetAsp 502
|||||
1517 TTGTTGTACTGAAAGAAATCTATGACTGAAAAAGAAAGTAAATGAT 1566
|||||

```

```

503 TyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyValAla 519
|||||
1567 TACGTGTCTACTCAAGTTTCAATGCAAAAGCTTGGGTGGTGGTCCG 1616
|||||
519 gpPheValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLys 536
|||||
1617 TTTTGTGACAGAGTACTTAAGGCTCAGCTGTAATAATTGACGCTAAG 1666
|||||
536 laIleArgGluIleLeuLysTyrProValAlaLysMet 548
|||||
1667 CAATTAGAGAAATCTGAGAAACCAAGTTGCTAAGATG 1704
|||||

```

seq_name: /SID2/gcgdata/geneseq/geneseqn/NA1996.DAT.AAT33850

seq_documentation_block:

ID AAT33850 standard; DNA: 2019 BP.

XX AAT33850:

DT 12-NOV-1996 (first entry)

DE DNA encoding streptavidin/luciferase fusion protein.

XX Streptavidin; luciferase; fusion protein;

KW Streptomycetes avidinil; luciola lateralis; firefly;

KM recombinant production; industry; ds.

XX Synthetic.

XX Key

FT CDS Location/Qualifiers

FT CDS 1..2019

FT CDS /*tag= a

PN JP07289264-A.

XX 07-NOV-1995.

XX 27-APR-1994; 94JP-0090275.

XX 27-APR-1994; 94JP-0090275.

XX (KIRK) KIRKMAN CORP.

PA WP1: 1996-015269/02.

DR P-PSDB: AAM04208.

XX New mutant streptavidin (SA) gene - fused to a firefly luciferase

PT gene, for the recombinant prepn. of a SA-FL fused protein

PI Example 1; Pages 5-6; 12pp; Japanese.

PS The present sequence encodes a streptavidin/luciferase fusion

CC protein, comprising the Streptomycetes avidinil streptavidin gene

CC and the luciola lateralis (firefly) luciferase gene. The fusion

CC protein can be prepd. by inserting the recombinant DNA encoding

CC it into a Escherichia species microbe, culturing the transformed

CC microbe in a medium and collecting the fusion protein from the

CC culture. The fusion protein can be used in industry.

CC

XX

Sequence 2019 BP: 590 A: 412 C: 467 G: 550 T: 0 other;

SO

alignment_scores:

Quality: 2801.00

Ratio: 5.130

Percent Similarity: 100.000

Percent Identity: 99.451

alignment_block:

US-09-581-241-4 x AAT33850 ..

Align seg 1/1 to: AAT33850 from: 1 to: 2019

3 AsmMetGluAsnAspGluAsnIleValTyrGlyProGluPProPheThrPr 19

1282	CTCGAATAAATATGATTTATCAAAATTTAGTTGAATTCATCTGGCGGAGC	1331
319	aProLeuSerLysGluIleGlyGluAlaValAlaArgArpPheasnLeuP	336
1332	ACCTTTATCTAAAGAAATTTGGTGAGCTGTTGCTAGACGTTTAAATTAC	1381
336	roGluValaArgGluGlyTyrGlyLeuThrGluThrSerAlaIleLe	352
1382	CGGGTGTTGCTCAAGGCTATGCTTTTACAGAAACAACTCTGCATTTAT	1431
353	lleThrProGluGlyAspAspLysProGluValaSerGlyLysValaLpr	369
1432	ATCAGACCGGGAAGCCATGTATTAACCAAGGCTCTTCCGCAAGTTGGCC	1481
369	oleuPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGlyP	386
1482	ATTATATTTAAAGCAAAAGTTATCATCTTGTACTTAAAAAAAATTTGGGCC	1531
386	roAsnArGATgAgGlyGluValCysValLysGlyProMetLeuMetLysGly	402
1532	CGAAGACAGCTGGAGAAAGTTTGTGTAAAGGCTCTTATGCTTAAGAAAGT	1581
403	TyrValaAspAsnProGluAlaThrArgGluIleIleAspGluGlyLyr	419
1582	TATGTGATATATCCAGAACAGCAACAGAGAAATCATGATGATGAAGAGTTG	1631
419	pleuHISThrGlyAspIleGlyTyrTyrAspGluGluLysHisPhePheI	436
1632	GTTGCACACAGGAGATATGGGATTAACGTAGTAAGAAAAACATTTCTTTA	1681
436	leValaAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGlnValPro	452
1682	TCGTGATGCTTTGAAGCTTTTATATCAAAATACAAAGATATATCAAGTACCA	1731
453	ProAlaGluLeuGluSerValleuLeuGlnHisProAsnIlePheAspAl	469
1732	CCGTGCGAATTAAGAACTCTTCTTTGGCAACATCCAAATATTTTGATGC	1781
469	aglyValAlaGlyValProAspProIleAlaGlyGluLeuProGluValaLay	486
1782	CGCGCGTGTGGGCTTCCAGATGCTATACTGGTGAGCTTCCGGAGCTG	1831
486	alValaValleuLysGlyLysSerMetThrGluLysGluValaMetasp	502
1832	TTGTGTGACTTAAAAAGAAATCTATCTACGTGAAGAAAAAGATTAATGAT	1881
503	TyrValaIaSerGluValSerAsnAlaLysArgLeuArgGlyValaLar	519
1882	TACGTTCGTAGTCACAGTTTCAATGCAAAACGTTTCCGTGGGGTCCG	1931
519	gPheValaAspGluValProLysGlyLeuThrGlyLysIleAspGlyLysa	536
1932	TTTTTGGGAGCAAGTAACTTAAGGCTCTCCTGTTAAATTTGACGGTAAG	1981
536	laIleArgGluIleLeuLysProValAlaLysMet	548
1982	CAATTAGAGAAATTAAGTAAGAAACCAAGTTGCTTAAGATG	2019
seq_name: /SIDS2/gcdata/geneseq/geneseqn/NA1996.DAT: AAT33851		
seq_documentation_block:		
ID	AAT33851 standard; DN: 2055 BP.	
AC	AAT33851:	
XX		
DT	12-NOV-1996 (first entry)	
XX		
DE	DNA encoding mutant streptavidin/Luciferase fusion protein.	
KM	Mutant: streptavidin; luciferase: fusion protein;	
KW	Streptomyces avidinii; Luciola lateralis; firefly;	
KW	recombinant production; industry; ds.	
XX		

419 pleuHisThrGlyAspIleGlyTyr-TyrAspGluGluLysHisPhePheI 436
 1668 GTTGACACAGAGATATTGGATTAGATGAGAAAGAAACATTCTTTA 1717
 436 IValAspArgLeuSerLeuIleLysTyrLysGlyTyrGlnValPro 452
 1718 TCGTGATCGTTTGAAGTCTTTAACTCAATCAAGGATATCAAGTACCA 1767
 453 ProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePheAspAl 469
 1768 CTGCTGATTAAGATCTGTTTGGACATCAATATTTTGGATGC 1817
 469 aGlyValAlaGlyValProAspProIleAlaGlyGluLeuProGlyAlaVal 486
 1818 CGGCTGCTGCGCTCCAGATCCTAAGCTGAGCTCCGAGACTG 1867
 486 AlValAlaLeuLysGlyLysSerMetThrGluLysGlnValMetAsp 502
 1868 TTGTTGCTACTGAAAGAAAGAAATCTATGACTGAAAGAAAGATATGAT 1917
 503 TyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyValAla 519
 1918 TACGTTGCTACTCAAGTTTCAATGCAAAAGCTTGGTGCTGTCGCG 1967
 519 gPheValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLysA 536
 1968 TTTTGTGACAGAGTACCTAAGGTCCTGCTGCTGCTGCTGCTGCTGCT 2017
 536 LaIleArgGluIleLeuLysLysProValAlaLysMet 548
 2018 CAATTAGAGAAATACGAGAAACCACTGCTGCTGCTGCTGCTGCTGCT 2055
 seq_name: /SDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV23580
 seq_documentation_block:
 ID AAV23580 standard; cDNA to mRNA; 2364 BP.
 AC AAV23580;
 XX
 DT 16-JUL-1998 (first entry)
 DE Antibody-firefly luciferase fusion protein gene.
 XX
 KW Firefly luciferase; antibody-luciferase fusion protein; ds.
 OS Luciola cruciata.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..2364
 FT /tag- a
 FT /transl_except- (pos: 670..672, aa: Gln)
 FT /transl_except- (pos: 739..741, aa: Trp)
 FT /transl_except- (pos: 1369..1371, aa: Ala)
 FT /note= "no stop codon given"
 XX
 PN JP09187281-A.
 XX
 PD 22-JUL-1997.
 XX
 PF 09-JAN-1996; 96JP-0001812.
 XX
 PR 09-JAN-1996; 96JP-0001812.
 XX
 PA (KIKK) KIKKOWAN CORP.
 XX
 DR WPI: 1998-275089/25.
 DR P-PSDB; AAM53882.
 XX
 PT Antibody-firefly luciferase fused protein - and related products
 PT i.e. firefly luciferase fused gene, recombinant DNA and its
 PT preparation
 XX

PS Disclosure: Page 13; 17pp; Japanese.
 XX
 CC This sequence encodes the fusion protein of the invention. The protein is
 CC a antibody-firefly luciferase fusion protein, in which an antibody part
 CC consisting of a peptide having antibody activity is combined with an
 CC enzyme part consisting of firefly luciferase.
 XX
 SQ Sequence 2364 BP; 697 A; 444 C; 559 G; 664 T; 0 other;
 alignment_scores:
 Quality: 2800.00 Length: 545
 Ratio: 5.138 Gaps: 0
 Percent similarity: 100.000 Percent identity: 99.633
 alignment_block:
 US-09-581-241-4 x AAV23580 ..
 Align seg 1/1 to: AAV23580 from: 1 to: 2364
 4 MetGluAsnAspGluAsnIleValTyrGlyProGluProPheTyrProI 20
 730 CTGAGAACGATGAAATTTTGTATGCTCTGAAACATTTTACCTAT 779
 20 eGluGluGlySerAlaGlyAlaGlnLeuArgLysTyrMetAspArgTyrA 37
 780 TGAAGAGGATCTGCTGAGACAAATGGCGCAAGTATATGATGATGATG 829
 37 IalysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGlyValAspTyr 53
 830 CAAACTTGGACCAATGCTTTACTTAACGCACTTACCGGGTGCATTAT 879
 54 ThrTyrAlaGluTyrLeuGluLysSerCysLeuGlyGlnAlaLeuLys 70
 880 ACGTAGCCGCAATFACTAGAAATATCATGCTGTCTAGAGAGAGCTTAA 929
 70 sAsnTyrGlyLeuValAlaAspGlyArgIleAlaLeuCysSerGluAsn 87
 930 GAATTATGGTTGGTTGTTGATGAGAAATGCGTTATGACATGAAACT 979
 87 yGluGluPhePheIleProValIleuAlaGlyLeuPheIleGlyValGly 103
 980 GTGAGAGTTCTTATCTCTGATTAAGCGGTTTATTAAGGTGTCGT 1029
 104 ValAlaProThrAsnGluIleTyrThrLeuArgGluLeuValHisSer 120
 1030 GTGCTCCCACTAATGAGATTACACTACGTGATGTTGTTACAGTTT 1079
 120 uGlyIleSerLysProThrIleValPheSerSerLysLysGlyLeuAsp 137
 1080 AGCATCTCTAAGCCACAAATGTTATGTTCTTAAAGAAAGATTAGATA 1129
 137 ySValIleThrValGlnLysThrValThrAlaIleLysThrIleValIle 153
 1130 AAGTTATTAAGTGTACAAAAAAGCGTAAGTCTATTAAACCATTTGTATA 1179
 154 LeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAspAsnIle 170
 1180 TTGACACGCAAGGATTAAGATTATCAATCATGAGCAACTTTAT 1229
 170 eLysLysAsnThrProGlnGlyPheLysGlySerSerPheLysThrValG 187
 1230 TAAAAAAAACACTCCACAAGTTTCAAGATCAAGTTTAAACCTGTAG 1279
 187 IuValAsnArgLysGluGlnValAlaLeuIleMetAsnSerSerGlySer 203
 1280 AAGTTAACCGCAAGAACCAAGTCTCTTATTAATGAACCTCTCGGGTTCA 1329
 204 ThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLeuValThr 220
 1330 ACCGGTTTGCCAAAAGGTGTCACACTTACATCATGAATTTGGTCACGCG 1379
 220 gPheSerHisAlaArgAspProIleTyrGlyAsnGlnValSerProGlyT 237

```

|||||
1380 TTTTTCACGGTAGAGATCAATTATGGAACCAAGTTTCACAGGCA 1429
237 hralalleuthrvala1prophenishisglypneclymethetr 253
|||||
1430 CGCGTATTTTAACTGATGACATTCACATCGTTTGTGATGTTTACT 1479
254 ThleuglytyrleuthrCysglsipheargilevalmetleuthrlysh 270
|||||
1480 ACTTTAGGCTATCTACTGTGGTTTGTGATGATGATGTAACGAAATT 1529
270 eaSPGIuThrPheLeuLysThrLeuGlnAspTyrLysCysSerSery 287
|||||
1530 TGACGAGAGACTTTTAAAAACACTGCAAGTTTACAATGTCTACAGCG 1579
287 a1lleuValProthrLeuPheAla1leuAsnArgSerGluLeu 303
|||||
1580 TTAATCTTGACCGACTTGTGCAATCTTAATGAAAGTGAATTACTC 1629
304 AspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGlyValApr 320
|||||
1630 GATTAATATGATTTATCAAAATTTAGTTGAATTCATCTGGCGGACACC 1679
320 oleuSerLysGluIleGlyGluAlaValAlaArgArgPheAsnLeuProG 337
|||||
1680 TTTATCTAAAGAAATGTGGAACCTGTGCTAGACGTTTAAATTTACCG 1729
337 lyValArgGlnGlyTyrGlyLeuThrGluThrSerAlaIleIleIle 353
|||||
1730 GTGTGCTCAAGGCTATGTTTAAAGAAACAACCTCTCAATTAATTATC 1779
354 ThrProGluLysAspLysProGlyAlaSerGlyLysValAlaProle 370
|||||
1780 ACACCGGAGGCGATGATTAACCAAGTGTCTGCGCAAGTTTGCCATT 1829
370 upheLysAlaLysVal1leAspLeuAspThrLysLysThrLeuGlyProA 387
|||||
1830 ATTTAAAGCAAAAGTATCGATCTGATATAAAAAAATTGGGCCCA 1879
387 snArgArgGlyGluValCysValLysGlyProMetLeuMetLysGlyTyr 403
|||||
1880 ACAGACGTGAGAAAGTTGTGTAAGGGTCTATAGCTTAAGAAAGGTAT 1929
404 ValAspAsnProGluAlaThrArgGluIle1leAspGluGluGlyTyr 420
|||||
1930 GTAGATTAATCCAGAGCAACAAGAAATATATGATGAACAAGCTTGCT 1979
420 uHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisPhePheIle 437
|||||
1980 GCACACAGAGATATTGGTATTACGATGAAGAAAAACATTTCTTATCG 2029
437 alaAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGlnValProPro 453
|||||
2030 TGGATCTGTTGAAGCTTTTATCAAAATACAAAGATATACCAAGTCC 2079
454 AlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePheAspAla 470
|||||
2080 GCTAATATTAGATCTGTTCTTTGCAACATCCAAATTTTGTATGCCG 2129
470 yValAlaGlyValProAspProIleAlaGlyGluLeuProGlyAlaVal 487
|||||
2130 CGTTGCTGGGCTTCCAGATCTATAGCTGTGAGCTTCGCGGAGCTGTG 2179
487 alaValLeuLysGlyLysSerMetThrGluLysGluValMetAspTyr 503
|||||
2180 TTGTACTTGAAAGAAATCTATGACTGAAAGAAAGAAAGTAAGGATTAC 2229
504 ValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyGlyValArgP 520
|||||
2230 GTTGTGATGTAAGTTTCAAAATGCAAAAGCTTTGCTGTGCTGCCGTT 2279
520 eValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLysAla 537
|||||

```

```

2280 TGTGACGAGTACCTAAAGGTCCTACTGTAAATTGACGCTAAACCA 2329
537 leArgGluIleLeuLysProValAlaLysMet 548
|||||
2330 TTAGAGAAATTTCTGAGAGAAACCACTGCTAAGATG 2364
seq_name: /SID2/gcgdata/geneseq/geneseqn/NA1999.DAT.AAX25717
seq_documentation block:
ID AAX25717 standard; cDNA to mRNA; 1656 BP.
XX
AC AAX25717:
XX
DT 21-MAY-1999 (first entry)
XX
DE Firefly luciferase gene #3.
XX
KW Bioluminescent protein; catalytic efficiency; stability; firefly;
KW luciferase; chimeric; luciola cruciata; luciola lateralis;
KW photinus pyralis; primer; PCR; amplification; ss.
XX
OS Chimeric - Luciola lateralis.
OS Chimeric - Photinus pyralis.
XX
WO9902697-A1.
XX
PD 21-JAN-1999.
XX
PF 30-JUN-1998; 98WO-JP02936.
XX
PR 08-JUL-1997; 97US-0051917.
XX
PA (KIKK ) KIKKOMAN CORP.
XX
PI Hirokawa K, Kajiyama N, Murakami S;
XX WPI: 1999-120898/10.
XX DR P-PSDB; AAW99366.
XX
PT New bioluminescent protein with improved properties - has greater
PT catalytic efficiency and stability and is obtained by modification
PT of natural precursors
XX
PS Example 3; Page 30-31; 53pp; Japanese.
XX
CC The invention relates to the generation of bioluminescent proteins with
CC improved catalytic efficiency and stability. The proteins are generated
CC by addition, deletion or substitution of amino acids from a precursor
CC protein e.g. firefly luciferases, or by constructing chimeric luciferase
CC proteins from luciferase genes from e.g. Luciola cruciata, L. lateralis
CC and Photinus pyralis. This sequence represents an example of a chimeric
CC luciferase gene of the invention.
XX
SQ Sequence 1656 BP; 529 A; 283 C; 366 G; 478 T; 0 other;

```

alignment_scores:

```

Quality: 2703.00 Length: 543
Ratio: 5.043 Gaps: 0
Percent Similarity: 98.711 Percent Identity: 96.317

```

```

alignment_block:
US-09-581-241-4 x AAX25717 ..

```

```

Align seg 1/1 to: AAX25717 from: 1 to: 1656

```

```

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17
|||||
1 ATGGAAGAACATGAGAGAACGATGAAATATTGTATGCTGACCAACTT 50
17 eTyrProIleGluGluGlySerAlaGlyAlaGlnLeuArgLysTyrMet 34
|||||
51 TTACCTATTGAAGAGGATCTGCTGAGACACAAATTGCCAAGTATATG 100

```

34 sparglyAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
 101 ATGATATGCAAAACTTGAGCAATTCGTTTACTTAACGCACTTACCGGT 150
 51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGly 67
 151 GTGATTTATACGTACGCCGAATCTTAGAAAAATCATCTGCTGTGAGGA 200
 67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
 201 GCGTTTAAAGAAATTCGTTGCTTGTGATGAGAAATTCGCTTATGCA 250
 84 ergLysAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
 251 GTGAAACCTGGAAGAATTCCTTATTCCTGATATAGCCGGTTATTTATA 300
 101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
 301 GGTGCGGTGGCTCCCACTAATGAGATTACACTTACGTAAGTAATGCT 350
 117 HisSerLeuGlyIleSerLysProThrIleValPheSerSerLysGly 134
 351 TCACAGTTTAGCAATCTCTAAGCCACAACTTGTATTTAGTTCTTAAACAG 400
 134 LysLeuAspLysValIleThrValGluLysThrValThrAlaIleLysThr 150
 401 GATTAGATAAAGTTATACCTACAAAAACGGTAACGTCTATTAAACCC 450
 151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAs 167
 451 ATTGTATATTGGACAGCAAGAGGATTATAGAGTTATCAATCATGCA 500
 167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheI 184
 501 CAACCTTATTAATAAAAAACACTCCACAAGGTTTCAAGATCAAGTTTAA 550
 184 YstThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
 551 AAACCTGTAAGAGTTAACCGCAAAAGACAAAGTGTCTATATATGAACTCT 600
 201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnIle 217
 601 TCGGGTTCAACCGGTTGCCAAAAGGTGTGCACTTACTCATGAAATTTT 650
 217 uValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
 651 GGTCACTAGATTTTCTCAGCTAGAGATCCAAATTTATGGAACCAAGTTT 700
 234 erProGlyThrAlaIleLeuThrValAlaProPheHisHisGlyPheGly 250
 701 CACAGGCGCGGCTATTTTAACTGTAGTACCATTCATCATGTTTGGT 750
 251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValIleMet 267
 751 ATGTTTACTTACTTATAGGCTATCTAATCTGCTTGTGATTTGCAATGTT 800
 267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
 801 AACGAAATTTTACGCAAGAGACTTTTAAAAACACTCCAAAGATTACAAAT 850
 284 YsserSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
 851 GTTCAGACCGTTATTTCTGTACCGACTTGTGTAATTCCTTAAAGAGT 900
 301 GluLeuLeuAspLysTyrAspLysSerAsnLeuValGluIleAlaSerGly 317
 901 GAATTAATCATTAATATGATTTATCAAAATTTAGTTGAATTTGATCTCG 950
 317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheA 334
 951 CCGAGACACCTTTATCTAAAGAAATTTGGTGAAGCTGTGTGAGAGCTTTTA 1000

334 snLeuProGlyValAlaArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
 1001 ATTTACCGGGGTGTTCTCAAGGCTATGTTTAAAGAAACCAACCTCTGCA 1050
 351 IleIleIleThrProGluLysAspLysProGlyAlaSerGlyLysVal 367
 1051 ATTATTAACACACCGCAAGGCGATGATTAACCAAGGTGCTTCGCAAACT 1100
 367 ValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
 1101 TGTGCTATTATTAAAGCAAAAGTTATGATCTTACTATAAAAACTT 1150
 384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
 1151 TGGCGCCGACAGACGCTGAGAGATTGTGTAAAGGCTCTTATGCTTATG 1200
 401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGlu 417
 1201 AAAGTTATGTAGATTAATCCAGAACCAAGAGAAATCATAGATGAGGA 1250
 417 uGlyThrLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
 1251 AGTTGTGTGACACAGAGATATTGGGTATTACATGATGAAGAAAAACTT 1300
 434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
 1301 TCTTATGCTGATGCTGTTTGAAGTCTTATCAATATCAAAAGATATGAG 1350
 451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
 1351 GTGGCCCCGCTGAATTGGAATTCATATTGTACAAACCCCAACATCTT 1400
 467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
 1401 CGACGCGGCGCTGCGAGCTTCTCCGACGATGACGCCGCTAATCTCCG 1450
 484 lValAlaValAlaLeuLysLysGlyLysSerMetThrGluLysGluVal 500
 1451 CCGCGGTGTTGTTTGGAGCAGCAAGAGATGACGAAAGAAAGAGATC 1500
 501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
 1501 GTGATTAACGTGCGCAGTCAAGTAACAACCGCAAAAGTTGCCGAGG 1550
 517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
 1551 AGTTGCTGTTGTGACGCAAGATACCGAAAGTCTTACCGGAAACTCGAGC 1600
 534 lLysAlaIleArgGluIleLeuLysLys 543
 1601 CAAAGAAATCAGAGATCCTCATTAAG 1629
 seq_name: /SID2/gcgdata/geneseq/geneseqn/AA1989.DAT:AA91170
 seq_documentation_block:
 ID AA91170 standard; DNA; 1644 BP.
 AC AA91170;
 XX
 DT 14-JUN-1990 (first entry)
 XX
 DE Recombinant luciferase gene.
 XX
 KW Luciferase; ATP assay; ss.
 XX
 OS Luciola cruciata.
 XX
 PN EP301541-A.
 XX
 PD 01-FEB-1989.
 XX
 PF 28-JUL-1988; 88BP-0112233.
 XX

PR 29-JUL-1987; 87JP-0187724.
PR 29-JUL-1987; 87JP-0187725.
PR 20-AUG-1987; 87JP-0205194.

(KIKK) KIKKOMAN CORP.

XX Masuda T, Tatsumi H, Nakano E;

XX WPI: 1989-033443/05.

DR P-PSDB: AAP94367.

XX Recombinant luciferase gene - used for efficiently producing luciferase
in *E. coli*, obtaining a stable and highly active prod.

PS Claim 2; fig 3; 35pp; English.

CC The sequence was obt'd. from cDNA prepd. from mRNA isolated from
CC *L. cricetata* tails. The DNA can be incorporated into vectors for
CC transformation of *E. coli* JM 1010.

XX Sequence 1644 BP; 529 A; 276 C; 338 G; 501 T; 0 other;

alignment_scores:

Quality: 2684.00 Length: 548
Ratio: 4.943 Gaps: 0
Percent Similarity: 99.088 Percent Identity: 93.431

alignment_block:

US-09-581-241-4 x AAN91170

Align seg 1/1 to: AAN91170 from: 1 to: 1644

```

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProH 17
1 ATGGAAACATGGAAGAGGATCTGCGAATTCATGATGACCTAACCGTT 50
17 eTyrProIleGluGlySerIleGlyAlaGluIleuArgLysTyrMet 34
51 TTACCCCTTCGGAAGAGGATCTGCGAATTCATGATGACCTAACCGTT 100
34 sPArGlyAlaLysLysGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 AGCGATATGCAAACTGGCGCAATTCCTTTACAATGCAATGCTACTGGT 150
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGly 67
151 GTTGATTAATCTTACCGCGAATTCCTGGAATAATCATGTTGTAGGAAA 200
67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
201 AGCTTTCGCAAAATTAATGCTTGGTTGGTATGCGAGAAATGCGTTATGA 250
84 eArgLysCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
251 GTGAAACTGTGAGAAATTTTATCTCTGTAATAGCCGCACTGTTATA 300
101 GlyValGlyValAlaProThrAsnGluIleGlyThrIleuArgLysLeu 117
301 GGTTAGGTGTTGCAACCACTAATGATTAATACCTTACCTGAACTGCT 350
117 IHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysLys 134
351 TCACAGTTTAGGTATCTCTAATACCAACATGTTAGTTCTTAAAG 400
134 ILeuAspLysValIleThrValGluLysTyrValIleAlaIleLysThr 150
401 GCTTAGATAAGTATATACAGTACAGAAACAGTAACCTATTAAACC 450
151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
451 ATTGTTATACATGATGACAAAGTTGATTAATGAGGATATCAATGCTGGA 500

```

```

167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPhe 184
501 CACCTTTATATAAAGAAACACTCCACAGGTTTCAAGATCCAGTTTCA 550
184 ySThrValGluValAsnArgLysGluGluValAlaLeuIleMetAsnSer 200
551 AAAGTGTGAAGTTGACCGTAAAGAACAACTGCTGCTATTAATGAACCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGlnSer 217
601 TCGGTTCTACCGGTTTGGCAAAAGGCGTACAACTTACTCAGCAAAATAC 650
217 uValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnVal 234
651 AGTCACATGATTTCTCATGCTAGAGATCCGATTTATGTAACCAAGTTT 700
234 eProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
701 CACGAGGACCGCTGTTTAACTGCTGCTCCATTCATCATGTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMet 267
751 ATGTCACTACTCTAGGATTAATTTGCTGTTTGTGCTGTTGATGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLys 284
801 AACAAATTCGATGAGAAACATTTTAAAACTCACAAGATTAATTAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTACAAAGTTATCTTGTACCGACCTGTTGCAATTCACAAACAAAGT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
901 GAATTAATCAATAATTAATGATTTGCAAAATTAATGATTAATGATTCG 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
951 CCGAGCACCTTATCAAAAGAAAGTTGTAAGCTGTTCTGACGCTTTA 1000
334 sPLeuProGlyValAlaArgGlnGlyTyrGlyLeuThrGlnThrSerAla 350
1001 ATCTTCCGCGGTGCTGCAAGGTTATGTTTACGAAACAAACATCTGCC 1050
351 IleIleIleThrProGluLysAspAspLysProGlyAlaSerGlyLys 367
1051 ATTATTAATTAACACCAAGAGACATTAACCAAGAGCTTCGGAAGT 1100
367 IValProLeuPheLysAlaLysValIleAspLeuAspThrLysThr 384
1101 CGTCCGCTGTTTAAAGCAAAAGTTATGATCTTGTACCAAAATCTT 1150
384 eArgLysProAsnArgArgGlyValAlaLysValLysGlyProMetLeu 400
1151 TAGGCTCAACAGCTGAGAAAGTTGTTAAAGCACTTATGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGln 417
1201 AAAGTTATGTAATAATATCCAGAACCAACAAAGAACTTATGACAGAA 1250
417 uGlyTyrPheLysThrGlyAspIleGlyTyrTyrAspGluLysHisP 434
1251 AGGTGGCTGCACACCGAGATATTTGGATTTATGTAAGAAAAACATT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
1301 TCTTTATGTCGATCGTTTGAAGCTTTTAAATCAATAACAAAGATACCA 1350
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIle 467
1351 GTACACACCTCCGAAATTAATCCGTTTCTTAAACATCCATCATCTCT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyLeuLeuPro 484

```

```

|||||
1401 TATATGTTGTTGCGCGGCTTCGATCCTGCTAGCTGCGAGCTTCAG 1450
484 1YAlaValAlaValLeuLysLysGlySerMetThrGluLysGluVal 500
1451 GAGCGGTTGTTGCTACTGAAACGGAATAATATGACCAAAAGAACTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgLysI 517
1501 AAGGATTTATGTTGCAAGTCAAGTTCAATGCAAAACGTTACGCTG 1550
517 YValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAsp 534
1551 TGTTCGTTTGTGATGAGTACCTTAAGAGTCTTACGGAATAATGACG 1600
534 1YValAlaIleArgGluLysLysProValAlaLysMet 548
1601 GCAGAGCAATTAGAGAAATCCTTAAGAACCAAGTTGCTAAGATG 1644
seq_name: /SDS2/gcgdata/geneseq/geneseq/NA1990.DAT.AA03801

```

seq_documentation_block:

ID AA03801 standard; cDNA; 1644 BP.

AC AA03801;

DT 22-AUG-1990 (first entry)
DT 26-FEB-1993 (revised entry)

DE Recombinant luciferase gene.

KW Luciferase; ATP; E. coli; photon; ds.

OS Synthetic.

PN JP02065780-A.

PD 06-MAR-1990.

PF 01-SEP-1988; 88bp-0216229.

PR 01-SEP-1988; 88bp-0216229.

PA (KIKK) KIRKMAN CORP.

DR WPI; 1990-113360/15.

XX P-PSDB; AAR05788.

XX Prepn. of luciferase -

XX using Escherichia sp. bacteria contg. recombinant DNA.

XX Claim 2; Page 416 + Fig 3; 20pp; Japanese.

XX Prepn. of luciferase comprises culturing Escherichia sp. contg.

XX recombinant DNA, and harvesting luciferase from the culture medium.

XX The recombinant DNA is a vector DNA into which has been inserted the

XX luciferase gene which has the nucleotide sequence below. Luciferase

XX can be produced quickly and efficiently with an increased photon

XX output, useful in assaying ATP.

XX Sequence 1644 BP; 529 A; 276 C; 338 G; 501 T; 0 other;

XX Alignment scores:

XX Quality: 2684.00

XX Ratio: 4.943

XX Percent Similarity: 99.088

XX Percent Identity: 93.431

XX Alignment_block:

XX US-09-581-241-4 x AA03801

XX Align seg 1/1 to: AA03801 from: 1 to: 1644

```

1 MetCysMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17
1 ATGGAAACATGGAAGAAACGATGAAATATGTGATGGACCTAACCGTT 50
17 eTyrProIleGluGlySerAlaGlyAlaGluIleuAsnArgLysTyrMe 34
51 TTACCCATGGAAGAGGATCTCTGGACACACATTAACCAATTAATG 100
34 sPArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 AGCGAATGCAAAATCTGGCGCAATTGCTTTTCAATATCAGTTACTGT 150
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGly 67
151 GTTGATTTATCTTACCGCCGAAATCTTGGAATAATCATGTTGCTAGAAA 200
67 uAlaLeuLysAsnTyrGlyLeuValValAspGluArgIleAlaLeuGly 84
201 AGCTTTGCAAAATATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 250
84 eTyrLysCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
251 GTGAAACATGTAAGAAATTTTATCTCTGTAATAGCCGACTGTTTATA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuAsnGluLeu 117
301 GGATGATGTTGTTGCAACCTTAATGATTAATGATTAATGATTAATG 350
117 HisSerLeuGlyIleSerLysProThrIleValPheSerLysLysG 134
351 TCACAGTTAGTATCTTAACCAACCAATCTATTTAGTTCTAAAGAAAG 400
134 1YLeuAspLysValIleThrValGluLysThrValAlaIleLysThr 150
401 GCTTAGATTAAGTTTAAACAGTACGAAACAGTAACTAATTAATTAAC 450
151 1YValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
451 ATTGTTATCTAGTATGCAAGATTGATTCAGAGATTCATCATCTGCA 500
167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerPheI 184
501 CACCTTTATTAAGAAACAACTCCACAGGTTTCAACAGATCAGATTGCA 550
184 ySthValGluValAsnArgLysGluGluValAlaIleLeuIleMetAs 200
551 AAACGTGGAAGTTGACCGCTAAGACAGATGCTCTTATATATATACCT 600
201 SerGlySerThrGlyLeuProLysGlyValGluLeuThrHisGluAs 217
601 TCGGTTCTACCGGTTTGCCTCAAGGCTACACTCTCAGAAATATAC 650
217 uValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGluVal 234
651 AGTCACTGATTTCTCATGCTAGATCCGATTTATGAGTAACCAATTT 700
234 eProGlyThrAlaIleLeuThrValAlaProPheHisIleGlyPheGly 250
701 CACAGGACCGCTGTTTAACTGCTCCATTCATCATCATGATGTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMet 267
751 ATGTTCACTACTAGGGTATTTAATTTGAGTTTCTGTTGTTGTTGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGluAspTyrLys 284
801 AACAAATTCGATGAAGAAACATTTTAAAACTCTACAAAGATTTAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTACAACTGTTATTTCTTGTACGACCTTTGTTGCAATTCACAAAG 900
301 GluLeuLysPlyTyrAspLeuSerAsnLeuValGluIleAlaSerG 317

```

|||||:::|||||
901 GAATTAAGTAAATACATTTGTCATTTAGTGAATTCATGACATCTGG 950
317 yglYalAProLeuSerLysGluIleGlyGluAlaIaArgArphea 334
|||||:::|||||
951 CGGAGCACCCTTATCAAAAGAGTTGTGAAGCTTGTCTAGACGCTTA 1000
334 snLeuProGluYalArgGlnGlyLeuThrGluThrSerAla 350
|||||
1001 ATCTTCCCGTGTGTCGAAGGTTATGTTTACAGAAACACATCTGCC 1050
351 llelleIleThrProGluLysAspLysProGlyAlaSerGlyLysVa 367
|||||
1051 ATTATTATTACACAGAGAGACGATTAACCGAGCTTCGGAAGT 1100
367 lVaIProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
|||||
1101 CGTGCCGTGTGTTAAAGCAAAAGTTATGATCTGTATACCAAAAATCTT 1150
384 euGIYProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
|||||
1151 TAGGTCCTACAGACGTCGAGAGATTGTGTTAAAGACCTATGCTTATG 1200
401 LysGlyTyrvAlaAspAsnProGluAlaThrArgGluIleIleAspGluI 417
|||||
1201 AAAGTTATGTAATATATCCAGAACACAAAGAACTTATTGACGAGAG 1250
417 uGIYTPLeuHisThrGlyAspIleGlyTyrvAspGluGluLysHisP 434
|||||
1251 AGGTGGCTGCACACCGAGATTTGATATGTATGTAAGAAACATTT 1300
434 hepheiLeuValAspArgLeuLysSerLeuIleLysTyrvLysGlyTyrvGln 450
|||||
1301 TCTTTATGTGATCGCTTGAGAGCTTTATATCAAAATACAAAGATACCA 1350
451 ValProProAlaGluLeuGluSerValIleuLeuGlnHisProAsnIleph 467
|||||
1351 GTACCACTGCGCAATTAGAATCCGTTCTTTGCACATCATCTATCTT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
|||||
1401 TGATGCTGGTGTGCGCGCTTCCTGATCCTGTAGCTGGGAGCTTCAG 1450
484 lYalAValIleuLysLysGlyLysSerMetThrGluLysGluVal 500
|||||
1451 GAGCCGTTGTTGTTACTGAAAGCGGAAAAATATGACCGAAAAAGAACTA 1500
501 MetAspTyrvAlaAlaSerGlnValSerAsnAlaLysArgLeuArgGlyG 517
|||||
1501 ATGATTTATGTTCAGAGTCAAGTTTCAATGCAAAACGTTTACGTGTGG 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
|||||
1551 TGTTCGTTTGTGTGATGAAGTACCTAAAGGTCTTACTGGAATAATTCAGC 1600
534 lYysAlaIleArgGluIleuLysLysProValAlaLysMet 548
|||||
1601 GCAGAGCAATTAGAGAAATCTTTAGAAACCAAGTGTCTAAGATG 1644

34 spArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
 101 ATGATATGCAAAACTGGACCAATGCTTTACTTAACGCACTTAACCGGT 150
 51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysLeuGlyGly 67
 151 GTGATATATAGTACGGCAATGCTTAAATGATGCTTCTTAAAGGAGA 200
 67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuGly 84
 201 GGCTTTAAAGATTAATGCTTGGTTGTTGATGAGAGAAATTCGTTATGCA 250
 84 ergLysAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
 251 GTGAAACTGAGAGAAATCTTATCTCTGATTAAGCCGGTTATTATATA 300
 101 GlyAlaGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
 301 GGATCGGTGGTGGCTCCAACTAAATGAGATTTACCTACGTGAATGGCT 350
 117 LHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysLysG 134
 351 TCACAGTTTGGGCACTCTTAAGCCAAACAAATTTGATTAGTTCTAATAAG 400
 134 LysLeuAspLysValIleThrValGluLysThrValThrAlaIleLysThr 150
 401 GATTAGATTAAGTTATTAACCTCAAAAAAAGCGTAACGTGCTATTAAGAAC 450
 151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetLys 167
 451 ATGTCTTATATGGACAGCAAGATGAGATTATAGAGTTATCAATCCATGAGA 500
 167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheIle 184
 501 CACTTATTAATAAAAAACACTCCACAAGTTTCAAAAGATCAAGTTTAA 550
 184 YThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
 551 AAACGTGAGAGTTAACCCGAAAGAACAAAGTGTCTTATTAATGAACCTCT 600
 201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnIle 217
 601 TCGGTTCAACCGGTTTGCACAAAAGGTGTCACACTTACTCATGAATATAT 650
 217 eValIThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
 651 CGTCACTAGATTTTCTCACGCTAGAGATCCAAATTTATGAAACCAAGTTT 700
 234 ePProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
 701 CACGAGGCAAGGCTATTTTAACGTGATGATCCATTCATGATGGTTTGGT 750
 251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
 751 AAGTTTACTACTTTAGGCTATCTAATCTGATGTTTGGTTTGTGTCATGTT 800
 267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
 801 AACGAAATTTGAGAGAGACCTTTTAAAAACACGCAAGATTACAAAT 850
 284 YSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
 851 GTTCAAGCGTTATCTTGTACCGACACTTGTTCGCAATTTCTTAATAGAAAGT 900
 301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGly 317
 901 GAATTAATGATAAATATGATTTATCAAAATTTAGTTGAATTCATCAGCTGG 950
 317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheA 334
 951 CGGAGCACCTTATCTTAAGAAATGTTGTAAGCTGTGCTAGACGTTTAA 1000

334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
 1001 ATTTACCGGGGCTTCTGTCAGGCTATGCTTTACAGAAACAACTCTGCA 1050
 351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVal 367
 1051 ATTATTAATCACACCGAAGCGGATGATTAACACAGTGCTTCTGCAAAAGT 1100
 367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysIleThrL 384
 1101 TGTCCCATTAATTTAAAGCAAAAGTTATCGATCTTGATCTAATAAAACTT 1150
 384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
 1151 TGGGCGGCAACAGACGTCGAGAACTTTGTGTAAGGGTCCATATGCTTATG 1200
 401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGluGly 417
 1201 AAAGTTATGTAGAAATCCAGAACCAAGAACAAATCATAGATGAAGA 1250
 417 uGlyThrLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
 1251 AGGTTGTTGCACACAGAGATATTTGGTATTAAGATGAAGAAAAACATT 1300
 434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
 1301 TCTTATCGTGAGATCGTTTGAAGCTTTAATCAATCAAAAGATATACAA 1350
 451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePhe 467
 1351 GTACCACTCTGTGAATTAAGATCTGTTCTTTGCAACATCCAAATATTTT 1400
 467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyLeuLeuProG 484
 1401 TGATCGCGCGGCTTCTGCTGCTCCAGATCTTATAGCTGGTGAAGCTTCGG 1450
 484 lValAlaValAlaLeuLysLysGlyLysSerMetThrGluLysGluVal 500
 1451 GAGCTGTTGTGTTACTTAAGAAAGAAATCTATGACTGAAGAAAGAGTA 1500
 501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyGly 517
 1501 ATGATTAATCGTTGCTAGTCAAGTTTCAAAATGCAAAACGTTTGGCTGTGG 1550
 517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
 1551 TGTCCGTTTGTGAGCAGAGTACCTAAAGGTCTCACAGTGAATAATTGACG 1600
 534 lYysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
 1601 GTAAAGCAATTTAGAGAAATCTGAAAGAACCAAGTGTGTAAGATG 1644
 seq_name: /SIDS2/gcdata/geneseq/geneseqn/NA1999.DAT:AAx84395
 seq_documentation_block:
 ID: AAx84395 standard; DNA; 1644 BP.
 AC: AAx84395;
 DT: 09-SEP-1999 (first entry)
 DE: L. lateralis luciferase coding sequence.
 XX: Luciferase; surfactant-tolerant; firefly; intracellular ATP assay; ss.
 OS: Luciola lateralis.
 XX: PN: W09933997-A1.
 XX: 08-JUL-1999.
 PD: 24-DEC-1998; 98WO-JP05864.
 PF: XX

PR 26-DEC-1997; 97JP-0361022.

XX (KIKK) KIKKOMAN CORP.

XX Hattori N, Murakami S;

XX WPI; 1999-419109/35.

DR P-PSDB; AAY22182.

XX Modified luciferase tolerant to surfactants and useful for assay of

PT intracellular ATP

XX Example 2; Page 34-39; 56pp: English.

XX This sequence encodes a luciola lateralis luciferase of the invention.
 CC The protein is a surfactant-tolerant luciferase which is derived from
 CC firefly luciferase by replacement of the glutamic acid residue at
 CC position 490 by another amino acid. The luciferase can be used in an
 CC assay of intracellular ATP in biological samples containing cells, for
 CC example for establishing the presence of cells, and their quantity, in
 CC the sample. The method can be applied to foodstuffs, drugs,
 CC agrochemicals, beverages and environmental samples. The activity of the
 CC surfactant-tolerant luciferase is not lowered by the presence of high
 CC concentrations of surfactant and therefore an assay of high accuracy can
 CC be carried out in a simple manner.

XX Sequence 1644 BP; 528 A; 261 C; 349 G; 506 T; 0 other;

alignment_scores:

Quality: 2821.00 Length: 548
 Ratio: 5.148 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.818

alignment_block:

US-09-581-241-6 x.AAX84395 ..

Align seg 1/1 to: AAX84395 from: 1 to: 1644

1 MetGluAsnMetGluAsnAspGluAsnIleValIYrGlyProGluProPh 17
 1 ATGGAAACATGAGAGACGATGAAATATGTTGATGCTGCTGAACCAT 50
 17 eYrProIleGluGluGlySerAlaGlyAlaGluLeuArgLysTyrMetA 34
 51 TTACCTATTTGAAGAGGATGCTGAGACACAAATTGGCAAGATATAGG 100
 34 SPArGTYrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
 101 ATCGAATATCAAAACTGAGCAATTGCTTTACTACGCACTTACCGGT 150
 51 ValAspTYrThrTYrAlaGluTYrLeuGluLysSerCysGlyLeuGly 67
 151 GTGCTATTATACGATACCGCAATCTTAGAAAAATCATGCTGCTAGGACA 200
 67 uAlaLeuLysAsnTYrGlyLeuValValAspGlyArgIleAlaLeuCys 84
 201 GGCTTTAAAGAAATTATGTTGGTTTGTATGAGACAAATTGGCTTAGCA 250
 84 eRGLuAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
 251 GTGAAAACGTGAAAGAAATCTTATTCCTGATTTACCGGTTATTATTA 300
 101 GlyValGlyValAlaProThrAsnGluIleTYrThrLeuArgGluLeuVal 117
 301 GGTCGTGGTGTGCTCCAACTATGAGATTACACTCTAGCGTAATTTGGT 350
 117 HisSerLeuGlyIleSerLysProThrIleValIlePheSerSerLysGly 134
 351 TCACAGTTTAGGCATCTCTAAGCAACAATTTGATTTAGTTCTAAAAAG 400
 134 LysLeuAspLysValIleThrValGluLysThrValThrAlaIleLysThr 150

401 GATTAGATAAAGTATTAAGTATACAAAAAAACGTAAGCTATTAATAAAC 450
 151 IleValIleLeuAspSerLysValAspTYrArgGlyTYrGlnSerMetAs 167
 451 ATGTGTTATATGACAGCAAAAGTGTATTAAGGTTATCAATTCATGCA 500
 167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerPheIle 184
 501 CAAGCTTTATTAATAAAACACCTCCACAAAGTTTCAAGAGTACAGTTTA 550
 184 ySThrValGluValAsnArgLysGluValAlaLeuIleMetLysSer 200
 551 AAAGCTAGAGATTAAACCGCAAAAGAAAGTGTCTCTTATTAATGAACCT 600
 201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnI 217
 601 TCGGGTTCAACCGCTTGGCCAAAAGGTGTGCACTTACATGAAATATT 650
 217 eValThrArgPheSerHisAlaArgAspProIleTYrGlyAsnGlnValS 234
 651 GGTCACTAGATTTTCTCAGCTAGACATCCAAATTATGGAACCAAGTTT 700
 234 eRProGlyThrAlaIleLeuThrValValProPheHisHsGlyPheGly 250
 701 CACAGGACAGCGCTATTTTAAGTGTAGTACCATTCATCATGATGTTGGT 750
 251 MetPheThrThrLeuGlyTYrLeuThrCysGlyPheArgIleValMetLe 267
 751 ATGTTTACTACTTATGAGCTATCTTAAGTGTGTTTTCGATTTGTCATGTT 800
 267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTYrLysC 284
 801 AACGAATTTGACAGCAAGACATTTTAAAAACATGCAAGATTACAAAT 850
 284 ySserSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
 851 GTTCACGCTTATCTTGTACCGACTTGTGTTGCAATTTCAATTTAATAGA 900
 301 GluLeuLeuAspLysTYrAspLysSerAsnLeuValGluIleAsnSerG 317
 901 GAATTACTCGATAATATGATTTATCAATTTAGTTGAATTCGATCTGG 950
 317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheA 334
 951 CGAGACACCTTATCTAAAGAAATGCTGAAGCTGTGCTAGACGTTTAA 1000
 334 snLeuProGlyValArgGlnGlyTYrGlyLeuThrGluThrSerAla 350
 1001 ATTTACCGGCTGTGTCACAGGCTATGAGTTTACAGAAACACCTCTGCA 1050
 351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
 1051 ATTTATATCACACCGGAAGGAGATGATTAACACGAGTCTCTGCAAAAGT 1100
 367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThr 384
 1101 TGTGCCATTATTTAAAGCAAAAGTTATCGATCTGATCAATAAAAAACTT 1150
 384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
 1151 TGGGCCCAACAGACGCTGAGAGAGTTTGTATTAAGGCTCATATGCTTAG 1200
 401 LysGlyTYrValAspAsnProGluAlaThrArgGluIleIleAspGluG 417
 1201 AAAGGTTATGTGATTAATCCAGAACCAACAGAAATCATATGATGAAGA 1250
 417 uGlyThrPheLysHsThrGlyAspIleGlyTYrTYrAspGluGluLysHsP 434
 1251 AGCTGTGTTGCACACAGAGATATTGGTATTTAGCATATAGAAAAACATT 1300
 434 hePheIleValAspArgLeuLysSerLeuIleLysTYrLysGlyTYrGln 450
 1301 TCTTTATCGTGAATGCTTTGAAGTCTTTAATCAATTAACAAAGGATATCA 1350

[illegible]

```

FT      CDS      1648..1908
FT      /tag= C
FT      /note= "encodes biotinylated peptide portion of fusion
FT      protein"
XX      JP08308578-A.
XX      26-NOV-1996.
XX      PD
XX      24-APR-1995; 95JP-0098857.
XX      PF
XX      14-MAR-1995; 95JP-0054625.
XX      PR 27-JUL-1994; 94JP-0193798.
XX      (KIKK ) KIKKOMAN CORP.
XX      WPI; 1997-059697/06.
XX      P-PSDB; AAM12395.
XX      DR
XX      PT Fusion protein comprising firefly luciferase and biotinylated
XX      PT peptide - useful in a bio-luminescent analytical method for
XX      PT quantifying ligands
XX      PS Example 7; Page 11-12; 13pp: Japanese.
XX      CC This is the nucleotide sequence encoding a novel fusion protein which
XX      CC comprises the firefly (Luciola lateralis) luciferase protein and an 87
XX      CC amino acid biotinylated peptide designated the E.coli biotin carboxy
XX      CC carrier protein (BCCP-87). The fusion gene was generated by firstly
XX      CC amplifying the sequence encoding the E.coli (BCCP-87) from the E.coli
XX      CC genome and inserting the resultant sequence into the 3' end of the
XX      CC firefly luciferase gene in plasmid pHLf230. The biotin peptide sequence
XX      CC replaces the C-terminal Met residue of the wild type luciferase sequence.
XX      CC The plasmid was then introduced into E.coli JM101 for production of the
XX      CC fusion protein. The novel protein can be used in bioluminescence assays
XX      CC to quantify luciferase ligands which may modulate binding of luciferase
XX      CC to its receptor.
XX      SQ Sequence 1908 BP; 598 A; 329 C; 426 G; 555 T; 0 other;

alignment_scores:
      Quality: 2812.00      Length: 547
      Ratio: 5.141      Gaps: 0
      Percent Similarity: 100.000      Percent Identity: 99.634

alignment_block:
US-09-581-241-6 x AAT63269 ..
Align seg 1/1 to: AAT63269 from: 1 to: 1908

```

```

251 GTGAAACCTGAGAGTCTTATTCCTGATATAGCCGGTTATTATATA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuAlaGluLeuVal 117
117 HisSerLeuGlyIleSerLysProThrIleValPheSerSerLysG 134
301 GGTCGCGGTGGCTCCACATGAGATTACCTACCTACGGAATTCGT 350
351 TCACAGTTAGGATCTCTTAACCCCAACATTTGATTTACTTAAAGAG 400
134 ILeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
401 GATTAGATTAAGATTATTAAGTACCAAAAAACCGTAATCTGATTAAC 450
151 ILeuAlIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
451 ATGTTATATATGACAGCAAGGAGTTATAGAGTTATCAATCCATGGA 500
167 PAspPheIleLysAsnThrProGlnGlyPheLysGlySerPheL 184
501 CAACCTTATTAAAAAACAACCTCCACAGTTTCAAGATCAAGTTTAA 550
184 ySThrValGluValAsnArgLysGluGlnValAlaLeuIleMetLysSer 200
551 AAACGTGAGAAATTAAACCGCAAGAAAGTTGCTCTTATTAATGAACCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsn 217
601 TCGGGTTCACCGCGTTTCCAAAAGGTGTGCACTTACATGAAAATTT 650
217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnVal 234
651 GTCACGCGCTTTTTCACGCTAGATCCAAATTTATGAAACCAAGTTT 700
234 erProGlyThrAlaIleLeuThrValValProPheHisGlyPheGly 250
701 CACCAGCAGCGCTATTATTAACGTAGTACCAATTCATCATGAGTTGG 750
251 MePheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
751 ATGTTTACTACTTATAGGCTATCTAATCTTGCTTGGTTTGGTATGAT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
801 AACCAAAATTGACGAGAGACATTTTAAAAACACGCAAGATTACAAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTTCAGCGCTATTCTGTACCGACTTGTGCAATTCCTTAATAGACGT 900
301 GluLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
901 GAATTACTCGATTAATATGATTTATCAAAATTTAGTTGAATTCGATCT 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
951 CGGAGCACTTATCTAAAGAAATTTGTGAAGCTGTGACGAGTTTAA 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
1001 ATTACCGGGGCTTCTGTCAGGCTATGTTTAAACAGAAACACCTCGCA 1050
351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
1051 ATTAATTATCACACCGGAAGGCGATGATTAACACAGTCTTCTGCAAG 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysThrL 384
1101 TGTGCAATTAATTAAGCAAAAGTTATCGATCTGATCAATTAAGAACT 1150
384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
1151 TGGGCCCAACAGACGTGAGACAGTTTGTGTAAAGGTCATGCTTATG 1200

```

```

401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleLeuGlu 417
    |||||
1201 AAGGTTATATGATATATCCAGAACACAGAGAAATCATAGATGAGA 1250
    |||||
417 uGlyTyrLeuHisThrGlyAspIleGlyTyrTyrAspGluIleHis 434
    |||||
1251 AGGTGGTGGTGCACACAGAGATATGGGTATTCAGATGAAGAAACAT 1300
    |||||
434 hehHeIleValAspArgIleLysSerLeuIleLysTyrGln 450
    |||||
1301 TCTTTATCGTGGATCGTTGAAGCTTTAATCAATAACAAAGATATCAA 1350
    |||||
451 ValProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
    |||||
1351 GTACCACTGCTGATTAAGATCTGTTCTTTCACACATCCAAATATTT 1400
    |||||
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyIleuProG 484
    |||||
1401 TGATGCCGGCGTGTGGCGTCCAGATCTATAGCTGGTGAACCTCCGG 1450
    |||||
484 LValAlaValAlaLeuLysGlyLysSerMetThrGluLysGluVal 500
    |||||
1451 GAGCTGTGTGTGACTTGAAGAAAAAGAAATCTATGACTGAAAAAGAGTA 1500
    |||||
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
    |||||
1501 AAGGATTACGTTGCTGATCAAGTTCAATGCAAAACGTTTGGTGTGG 1550
    |||||
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAsp 534
    |||||
1551 TCTCGTTTGTGGAGAAAGTACCTAAAGGCTCACTGGTAAATTTGAGC 1600
    |||||
534 LysAlaIleArgGluIleLeuLysLysProValAlaLys 547
    |||||
1601 GTAAGCAATTAGAGAAATPACTGAAGAAACAGTTGCTAAG 1641
    |||||
seq_name: /SIDS2/gcgdata/geneseq/geneseq/NM198.DAT:AAV32467
seq_documentation_block:
ID   AAV32467 standard; DNA; 1920 BP.
XX
AC   AAV32467;
XX
DT   10-SEP-1998 (first entry)
XX
DE   Luciferase-lysostaphin fused gene.
XX
KW   Luciferase-lysostaphin fused gene; chimeric; bioluminescence; ds.
XX
OS   Chimeric - Staphylococcus simulans.
XX
CH   Chimeric - Luciola lateralis.
XX
FH   Key
FT   CDS
    1..1920
    /product= "Luciferase-lysostaphin fusion protein"
    /note= "Nucleotides 1-1647 encode luciferase
    protein while nucleotides 1648-1920 encode
    lysostaphin; CDS does not contain stop codon"
JP10150991-A.
XX
PD   09-JUN-1998.
XX
PF   25-NOV-1996; 96JP-0328042.
XX
PR   25-NOV-1996; 96JP-0328042.
XX
PA   (KIKK ) KIKKOMAN CORP.
XX
DR   WPI; 1998-379994/33.
DR   P-PSDB; AAW48856.

```

```

XX
PT   New Luciferase-lysostaphin fused protein - useful for
PT   bioluminescence analysis
XX
XX
PS   Disclosure; Pages 6-7; 10pp: Japanese.
XX
CC   The invention claims for a luciferase-lysostaphin fused gene which
CC   encodes a fused protein in which a peptide part consisting of
CC   luciferase, from Luciola lateralis, is connected to a peptide part
CC   consisting of lysostaphin from staphylococcus simulans. The method
CC   described by the invention can be used to prepare a luciferase-
CC   lysostaphin fused protein efficiently. The fused protein is useful
CC   for bioluminescence analysis.
XX
SQ   Sequence 1920 BP; 627 A; 310 C; 407 G; 576 T; 0 other;

alignment_scores:
    Quality: 2812.00      Length: 547
    Ratio: 5.141         Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 99.634

alignment_block:
US-09-581-241-6 x AAV32467 ..

Align seg 1/1 to: AAV32467 from: 1 to: 1920

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluPro 17
    |||||
1 ATGGAAGAACATGGAGAACGATGAAATATTTGTATGATGCTTAACCAT 50
    |||||
17 eTyrProIleGluGluGlySerAlaGlyAlaGlnLeuArgLysTyrMet 34
    |||||
51 TTACCTATTGGAAGAGGATCTGCTGACACACATTCGCCAGATATGG 100
    |||||
34 sPArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
    |||||
101 ATCGATATGCAAAACTGTGAGCAATGCTTTACTTAACGACACTTACCG 150
    |||||
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysGlyLeuGly 67
    |||||
151 GTGCAATTATACGACCGCGCAATCTTAGAAAAATCTGCTGTAGACA 200
    |||||
67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
    |||||
201 GCGTTTAAAGAAATTAAGTTGTTGTTGATGAGAAATGGGTATGCA 250
    |||||
84 eGluAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
    |||||
251 GTGAAAACTGTGAGAGATCTTATTCGTATTAGCCGTTTATTTATA 300
    |||||
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeu 117
    |||||
301 GGTGTCGGTGGTCCCAACTAATAGATTTACACTTACGTAATGGT 350
    |||||
117 HisSerLeuGlyIleSerLysProThrIleValPheSerSerLysG 134
    |||||
351 TCACAGTTAGGCATCTCTAACCCAATGTATTTAGTTCTTAAAAAG 400
    |||||
134 LysLeuAspLysValIleThrValGlnLysThrValThrAlaIleLys 150
    |||||
401 GATTAGATTAAGTTATTAACGTACAAAAACGGTAACTGATTAAAAAC 450
    |||||
151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
    |||||
451 ATTGTTATATTGGACAGCAAGTGAATTATAGAGTTTCAATTCATGGA 500
    |||||
167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPhe 184
    |||||
501 CAACCTTTATTAAAAAAACACTCCACAAGGTTTCAAGATCAATCACTTT 550
    |||||
184 yThrValGluValAsnArgLysGluGlnValAlaLeuIleMetLAsnSer 200
    |||||

```

551 AAAGTGTAGAGTTAAACCCGAAGAACAAAGTGTCTATATAGACTCT 600
 201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGlnAsnL 217
 601 TCGGGTTCACACCGGTTGGCCAAAAGGTGTCAACTTACTCATGAAATTT 650
 217 eValThrArgPheSerHisAlaArgAspProLysGlyAsnGlnValS 234
 651 GGTCACCGCGTTTCTCCAGCTAGAGATCCCAATTTATGAAACCAAGTTT 700
 234 ePProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
 701 CACACGACGACGGCTATTTTAACGTACTGATACCATTCATCGTTTGGT 750
 251 MetPheThrThrLeuGlyIleValThrCysGlyPheArgGlyIleValMet 267
 751 ATGTTTACTACTTATAGCTATCTAACTGGTGGTTTGTGATGCAAGTT 800
 267 uThrLysPheAspGlnGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
 801 AACGAAATTTGACGACAGACTTTTAAAAACACTGCACAGATACCAAT 850
 284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
 851 GTTCACGCGTATTTCTGTACCGACTTGTTCGAATTTCTTAATAGAGT 900
 301 GluLeuLeuAspLysIleValPheLeuSerAsnLeuValGlnIleAlaSerG 317
 901 GAATTTACTCGTAAATATGATTTATCAAAATTTAGTTAAATTCATCTGG 950
 317 yGlyAlaProLeuSerLysGlnIleGlyValAlaValAlaArgArgPhe 334
 951 CGGAGCAGCTTATCTAAAGAAATTTGGTGAAGCTGTGTGACAGCTTTTA 1000
 334 snLeuProGlyValArgGlnGlyIleValThrGlyLeuThrGlyThrSerAla 350
 1001 ATTTACCGCGGTTCGTCAAGCTATGTTTAAAGAAACCACTCTGCA 1050
 351 IleIleIleThrProGlnGlyAspAspLysProGlyAlaSerGlyLysVa 367
 1051 ATTATTTATCAACCGGAGGCGATGATTAACCAAGTCTTGTGGCAAAAGT 1100
 367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
 1101 TGTGCATTATTTAAAGCAAAAGTTATCATCTGATCTAATAAAACTT 1150
 384 euGlyProAsnArgArgGlyValValCysValLysGlyProMetLeuMet 400
 1151 TGGGCCCGACAGACGTGAGAGAGTTTGTGAAGGTCCTATGCTATATG 1200
 401 LysGlyTyrValAlaAspAsnProGlnAlaThrArgGlnIleIleAspGlnG 417
 1201 AAAAGTTATGTAGATATCCAGAACCAACAGAAATCATAGATGAGA 1250
 417 uGlyTyrPheHisThrGlyAspIleGlyTyrTyrAspGlnGlyLysHisP 434
 1251 AGCTGTGTTGCACACAGAGATATGGGTATTAACGATCAACAAAAACAATT 1300
 434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
 1301 TCTTTATCGTGGATGCTTGAAGTCTTTATCAAAATACAAAGGATATCA 1350
 451 ValProProAlaGluLeuGlnSerValLeuLeuGlnHisProAsnIlePh 467
 1351 GTACACACCGTGAATTAATCTGTTCTTTGCAACAATCCAAATATTTT 1400
 467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
 1401 TGTATGCCGCGCTGTGCGGTTCAGATCCATACCTGTTGAGCTTCCGG 1450
 484 lValValValValLeuLysGlyLysSerMetThrGlnLysGlnVal 500
 1451 GACGCTGTGTGTACTGAAAAAGGAAATCTATGACTGAAAAAGAGTA 1500

501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
 1501 ATGGAATTACGTTGCTAGTCAACTTTCAATGCAAAACGTTGCGTGG 1550
 517 yValArgPheValAspGlnValProLysGlyLeuThrGlyLysIleAspG 534
 1551 TGTCCGTTTGTGAGAGAGTACTTAAAGTCTCACTGGTAAATTTGACG 1600
 534 lLysAlaIleArgGlnIleLeuLysLysProValAlaLys 547
 1601 GTAAGCAATTTAGAGAAATATCTGAAGAAACCAAGTTGCTAAG 1641
 seq_name: /SID52/gcgdata/geneseq/NA1990.DAT:AA003257

seq_documentation_block:

ID AA003257 standard; DNA; 1644 BP.

AA003257:

22-JUL-1990 (first entry)

Luciola lateralis recombinant luciferase gene.

Luciferase; enzyme; ATP; plasmid pH1.7; firefly; ss.

Luciola lateralis.

Luciola lateralis.

Key Location/Qualifiers

CDS 1..1644

FT /*tag= a

FT /product=luciferase

FT /EC_number=EC-1.13.12.7

PR 22-DEC-1988; 88JP-0162402.

PR 01-JUL-1988; 88JP-0162402.

XX (KIRK) KIKKOMAN CORP.

XX PI Tatsumi H, KajiYama N, Nakano E;

XX DR WPI: 1990-038240/06.

XX DR P-PSDB; AAR03731.

XX PT New gene encoding luciferase

XX PT derived from Luciola lateralis and expressed in and purified

XX PT from Escherichia coli genus cells.

XX PS Claim 3; page 24; 43pp; English.

XX CC This enzyme is useful for determining ATP levels. It is

XX CC contained in plasmid pH1.7.

XX SQ Sequence 1644 BP; 529 A; 264 C; 348 G; 503 T; 0 other;

alignment_scores: quality: 2811.00 length: 548

Percent Similarity: 99.818 Percent Identity: 99.453

alignment_block:

US-09-581-241-6 x AA003257 ..

Align seg 1/1 to: AA003257 from: 1 to: 1644

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17

```

|||||
1  AAGGAAAAATGGAAGACGATGAAAAATATGTGATGTCCTGACCAT 50
17  eYrPrProlleuglucylSerAlaGluAlaGlnLeuAlaArgLysTyrMet 34
51  TTAACCTATGTAAGAGAGATGCTGAGACACAATGGCCAAAGTAATAG 100
34  sPaRgtYrAlaLysLeuGlyAlaAlaIleAlaPheThrAsnAlaLeuThrGly 50
101  ATCGATACGCCAAACTTGAGCAATGGCTTTTACTACCGACTTACC 150
51  ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGly 67
151  GTCGATTAACGACGCGCAATACCTAGAAAAATCATGCTGCTAGAGAG 200
67  uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
201  GGCTTAAAGAAATATGTTGGTTGTTGATGAGAAATGGCTTATGCA 250
84  eRgluAsnCysgluGluPhePheIleProValIleuAlaGlyLeuPheIle 100
251  GTGAAAACTGTGAGAAATCTTATCTGTAATACCGCTTATATTATA 300
101  GlyValGlyValAlaProThrAsnGluIleTyrThrLeuAlaGlyLeuVal 117
301  GGTCGCGGTGCTCCACACTATGAGATTACACTCTACGTAATGTGT 350
117  hHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysGly 134
351  TCACAGTTAGCATCTCTAAGCCAAATGTGATTTACTTCTAAAAAAG 400
134  lYleuAspLysValIleThrValGlnLysThrValIleAlaIleLysThr 150
401  GATTGATTAAGTTATTAAGTACAAAAAAGCGTAACGCTATTAAAC 450
151  lIleValIleLeuAspSerLysValAspTyrArgIleTyrGlnSerMetAs 167
451  ATTGTTATATGAGACGCAAAAGTATTAAGGTATCAATCCATGCA 500
167  PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheL 184
501  CAACCTTTATTAATAAAACACTCCACAAAGGTTTCAAGGATCAAGTTTA 550
184  ySThrValGluValAsnArgLysGluGlnValAlaIleuIleMetAsnSer 200
551  AAACGTGTAAGATTAAACCCGCAAGAACAGTGTCTTATTAATGAACCT 600
201  SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnIle 217
601  TCGGTTCAACCGGTTTGC AAAAGGTGCAACTTACTCATGAATAATGC 650
217  eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
651  AGTCACTAAGATTTTCTACGCTAGAGATCAATTTATGAAACCAAGTTT 700
234  eRProGlyThrAlaIleLeuThrValValProPheHisGlyPheGly 250
701  CACGAGGACGCGTATTTTAACTGTAAGCAATCCATCATGCTTGTGCT 750
251  MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
751  ATGTTTACTACTTAAAGCTATCTAATCTGTTGTTTCTGTAATGCATGTT 800
267  uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
801  AACGAAATTTGACGAGAGACTTTTAAACACGTCAGAAATTAACAAT 850
284  ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851  GTTCAACGCTATATCTGTACGACTTGTTCGATTCCTTAATGAAGT 900
301  GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGly 317
|||||

```

```

901  GAATTAATCGATAAATATGATTTATCAAAATTTAGTTGAATTCACATCTGC 950
317  yGlyAlaProLeuSerLysGluIleGlyAlaValAlaArgArgPhe 334
951  CGGAGCACCTTTATCTAAAGAAATGCTGAAGCTGTGCTAGACGTTTTA 1000
334  snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
1001  ATTTACCGGGGTGTTGCTCAAGGCTATGTTTAAACAAACACCTCTGCA 1050
351  lIleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
1051  ATTATATATACACCGGAGCGCATGATAACCGAGTCTTCTGCAAAAGT 1100
367  lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
1101  TGTGCCATTTATTTAAAGCAAAAGTTATCGATCTTGTACTTAAAAAACTT 1150
384  euGlyProAsnArgArgGlyGlyValAlaCysValLysGlyProMetLeuMet 400
1151  TGGGCCCGAAGACAGACTGCGAAGAGTTGTGTAAAGGCTCTTATGTTATG 1200
401  LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGlu 417
1201  AAAGTTATGTAGATTAATCCAGAACCAACAGAGAAATCATAGATGATAGA 1250
417  uGlyThrLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
1251  AGSTTGTTGACACAGAGATATGGAATGATTAACGATGAAGAAAAACATT 1300
434  hPheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
1301  TCTTATATCTGATGCTTGAAGTCTTTAATCAAAATACAAAGATATCA 1350
451  ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
1351  GTACCACTGCTGAATTAAGAACTGTTCTTTTGC AAACATCAAAATATTTT 1400
467  eAspAlaGlyValAlaGlyValAlaProAspProIleAlaGlyGluLeuProG 484
1401  TGATGCCGCGCTGCTGGGCTCCAGATCCCTATAGCTGGTGAAGTCCCG 1450
484  lYAlaValAlaValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
1451  GAGCTGTTCTGTTACTTGAAAAAGAAATCTATGACGAAAAAAGAGTA 1500
501  MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
1501  ATGGAATTACGTTGCTAGTCACAGTTTCAAAATGCCAAACGTTGCGGTGG 1550
517  yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
1551  TGTCCGTTTGTGGAGCAATCTAAGAGTCTCACTGCTGTAATAATGACG 1600
534  lYlYsAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
1601  GTAAAGCAATTAAGAAATTAAGAAACCAAGTGTGTAAGATG 1644
seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:AA763268
seq_documentation block:
ID  AA763268 standard; DNA; 1704 BP.
XX  AAT63268;
AC  AAT63268;
DT  20-MAY-1997 (first entry)
DE  Firefly luciferase conjugated with 23 aa peptide at 5' end.
XX  Fusion protein; firefly; luciola lateralis; luciferase; biotinylation;
KW  wild type; E.coli; bioluminescence assay; quantification; ligand;
receptor; ds.
XX

```



```

1167 ATTATTAAAGCAAAAGTTATCGATCTTGATCTTAAAGAAAGCTTTGGCC 1216
386 rosnrtrgarglglygluvalcysvallysglyprometleumetlysgly 402
1217 CGAAGACAGCTGGAGAGATTGTGTAAAGGGCTTATGCTTATGAAAGGT 1266
403 TyrValAspAsnProGluValThrArgGluIleLeuAspGluGlyTr 419
1267 TATGTAGATATATCCAGACAGCAAGCAAGAAATCATAGATCAAGAGCTTG 1316
419 pleuHstHnglyAspIleGlyTrTyrAspGluGlyShisPhePheI 436
1317 GTTGCAACAGAGGATTTGGTATTACGATGAAGAAACAAATTTCTTTA 1366
436 leValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGluValPro 452
1367 TCGTGATCGTTTGAAGTCTTTATCATCAATACAAAGATATCAAGTACCA 1416
453 ProAlaGluLeuGluSerValLeuLeuGluHisProAsnIlePheAspAl 469
1417 CCGCTGAATGAAATCTGCTTTTGCACCAATCCAAATATTTTGTATGC 1466
469 aGlyValAlaGlyValProAspProIleAlaGlyGluLeuProGlyValay 486
1467 CGGCGTGTGGGCTTCCAGATCTTATAGCTGGTGAAGCTTCCGGAGCTG 1516
486 aValValLeuLysLysGlyLysSerMetThrGluLysGluValMetAsp 502
1517 TTGTTGTAATGAAAAAGAAATATCATGACGAAAAAGAGTAATGAT 1566
503 TyrValAlaSerGluValSerAsnAlaLysArgLeuArgGlyGlyValar 519
1567 TACGTTGCTAGTCAGATTCAATGCAAAACGTTTGGCTGGTGGTCCG 1616
519 gPheValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLysA 536
1617 TTTTGTGAGCAAGTACCTAAAGGTCTCAGTAAATTTGACGGGTAAG 1666
536 laIleArgGluIleLeuLysLysProValAlaLysMet 548
1667 CAATTAGAGAAATACTGAAGAAACGTTGCTTACGATG 1704

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:AAV23595
seq_documentation_block:
ID AAV23595 standard; cDNA to mRNA; 1704 BP.
XX
AC AAV23595;
XX
DT 16-JUL-1998 (first entry)
XX
DE Antibody-Firefly Luciferase fusion protein gene.
XX
KM Firefly Luciferase; antibody-Luciferase fusion protein; ds.
XX
XX Luciola cruciata.
XX
OS
XX
FH Key location/Qualifiers
FT CDS 1..1704
FT /*tag= a
FT /note= "no stop codon given"
XX
XX JP09187281-A.
XX
XX 22-JUL-1997.
XX
XX 09-JAN-1996; 96JP-0001812.
XX
XX 09-JAN-1996; 96JP-0001812.
XX
XX (KIKK ) KIKKOMAN CORP.
XX
XX WPI; 1998-275089/25.

```

```

DR P-PSDB; AAV53883.
XX
XX Antibody-firefly luciferase fused protein - and related products
PT i.e. firefly luciferase fused gene, recombinant DNA and its
PT preparation
XX
XX PS Disclosure: Page 10-11; 17pp; Japanese.
XX
XX This sequence encodes a fusion protein of the invention. The protein is
CC a antibody-firefly luciferase fusion protein, in which an antibody part
CC consisting of a peptide having antibody activity is combined with an
CC enzyme part consisting of firefly luciferase.
XX
SQ Sequence 1704 BP; 541 A; 275 C; 361 G; 527 T; 0 other;

alignment_scores:
Quality: 2799.00 Length: 546
Ratio: 5.126 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.267

alignment_block:
US-09-581-241-6 x AAV23595 ..
Align seg 1/1 to: AAV23595 from: 1 to: 1704

3 AsnMetGluAsnAspGluAsnIleValIleTyrGlyProGluProPheTyrPr 19
:.....:
67 AGCTCGAGACAGATGAAAAATATGTGTATGCTCGAAGCAATTTTACC 116
19 oileGluGluGlySerAlaGlyValAlaGluLeuArgLysTyrMetAspArgT 36
117 TATGTAAGAGGGATCGCTGGAGACAAATTCGCGCAAGTATATGATGATGAT 166
36 yAlAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGlyValaAsp 52
167 ATGCAAAACCTTGAGCAATTCCTTTACTACGACCACTTACCGGTGTCGAT 216
53 TyrThrTyrAlaGluIleTyrLeuGluLysSerCysLysGluGluAlaIle 69
217 TATACGACCGCGAATACTTACAAAAATCATGCTGTCTAGAGAGGCTTT 266
69 uLysAsnTyrGlyLeuValAlaAspGlyArgIleAlaLeuLysSerGluA 86
267 AAGAAATTATGCTTGTGTTGATGAGAAAGATTCGTTATATGACAGGAAA 316
86 snCysGluGluPhePheIleProValIleAlaGlyLeuPheIleGlyVal 102
317 ACTGTGAAGAGTCTTATCTCTGATTTAGCCGGTTTATTTATAGGTGTC 366
103 GlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuValHisSe 119
367 GGTGTGGCTCCCAACTAATGAGATTACACTGATGATTTGTTCCACAG 416
119 rLeuGlyIleSerLysProThrIleValIlePheSerSerLysGlyLeuA 136
417 TTTAGGCAATCTCTTAAGCCAAATGTATTTAGTTCTAAAAAAGGATTAG 466
136 sPlyValIleThrValGluLysThrValIleAlaIleLysThrIleVal 152
467 ATAAAGTATATACTTACAAAAAAGGTAAGCTTATTTAAACCACTTGT 516
153 IleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAspAsnph 169
517 ATATTGACAGCAAGAGTGAATTAAGAGTTATCAATCCATGACAACTT 566
169 eileLysLysAsnThrProGluGlyPheLysGlySerSerPheLysThrV 186
567 TATTTAAAAAAACACTCCACAAGGTTTCAAGAGATCAAGTTTAAACTG 616
186 aGluValAsnArgLysGluGluValAlaLeuIleMetAsnSerSerSery 202
617 TAGAAGTTAACCGCAAGACAGATGCTCTTATATATGAACTCTTCGGGT 666

```

203 SerThrGlyLeuProIySGlyValGlnLeuThrHisGluAsnIleValIh 219
 |||
 667 TCACCGGTTGGCCAAAAGGTGGTCACTTACATGATAATTTGGCCAC 716
 219 rArgPheSerHisIaIarGspProIleTyrgIyAsnGlnValSerPro 236
 |||
 717 GCGTTTTCACGCTAGACATCCAAATTTATGGAACCAAGTTTCACGAC 766
 236 lYThrAlaIleLeuThrValValProPheHisHisGlyPheGlyMetPhe 252
 |||
 767 GCAGGCTATTTTACTAGTACATTCATTCATCATGTTTGGATGTTT 816
 253 ThrThrLeuGlyTyrlLeuThrCysGlyPheArgIleValMetLeuThrly 269
 |||
 817 ACTACTTAGGCTATCTAACTGTGGTTTCTGTAATGCTCATGTTAACGAA 866
 269 sPheAspGlnGluThrPheLeuLysThrLeuGlnAspTyrlLysCysSers 286
 |||
 867 ATTTGACGAGAGACTTTTAAACACCTGCAAGATTACAAATGTTCAA 916
 286 eValIleLeuValProThrLeuPheAlaIleLeuAsnArgSerGluLeu 302
 |||
 917 GCGTATTTCTGTACCGACTTGTGTTGCAATTTCTTAATAGAAAGTAATTA 966
 303 LeuAspLysTyrlAspLeuSerAsnLeuValGluIleAlaSerGlyGlyAl 319
 |||
 967 CTGCTAAATATGATTTATCAAAATTTAGTTGAAATTCATCTGGCGGAGC 1016
 319 aProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheAsnLeuP 336
 |||
 1017 ACCTTTATCTAAAGAAATTTGGTGAAGCTGTGCTAGACGTTTAAATTAC 1066
 336 rGlyValAlaArgGlnGlyTyrlGlyLeuThrGluThrSerAlaIleIle 352
 |||
 1067 CCGGTTGCTGCAAGGCTATGTTTAAACAGAAACCTCTGCAATTAATTT 1116
 353 lIethrProGluGlyAspAspLysProGlyAlaSerGlyLysValValPr 369
 |||
 1117 ATCAACCGGAGGCGATATTAACCAAGTCTCTGCGCAAAAGTTGTGCC 1166
 369 oLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGlyP 386
 |||
 1167 ATTATTTTAAAGCAAAAGTTATCGATCTTGATCTAAATAAACTTGGGCC 1216
 386 rGAsnArgArgGlyGlyValLysValLysGlyPrometLeuMetLysGly 402
 |||
 1217 CGAACAGACGTGAGAGAGTTTGTGTAAGGCTCTATCTTATGAAAGCT 1266
 403 TyrlValAspAsnProGluAlaThrArgIleIleAspGluGluGlyTr 419
 |||
 1267 TATGTAGATTAATCCAGAAAGCAACAGAAATCATATGATGAAGAGGTTG 1316
 419 pLeuHisThrGlyAspIleGlyTyrlAspGluGluLysHisPhePheI 436
 |||
 1317 GTTGCACACAGAGATATTTGGTATTCATGATGAAGAAAAACTTCTTTA 1366
 436 lValAlaAspArgLeuLysSerLeuIleLysTyrlGlyGlyTrGlnValPro 452
 |||
 1367 TCGTGATGCTTTGAAGCTTTTAAATCAATACAAAGGATATCAAGTATCA 1416
 453 ProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePheAspAl 469
 |||
 1417 CCTGCTGATTTAGAAATCTGTTCTTTGGCAACATCCAAATATTTTGATGC 1466
 469 aGlyValAlaGlyValProAspProIleAlaGlyGluLeuProGlyAlaVal 486
 |||
 1467 CCGCGTTGCTGGGCTCCAGATCTATACCTGCTGAGCTTCGGGAGCTG 1516
 486 aValValLeuLeuLysGlyLysSerMetThrGluLysGluValMetAsp 502
 |||
 1517 TTGTTGTACTGAAAAAGAAAAATCTATGACTGAAAAAGAAAGTAATGAT 1566

503 TyrlValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyGlyValAr 519
 |||
 1567 TACGTTGCTAGTCACTTCAATGCAAAACGTTTCCGGTGGTGTCCG 1616
 519 gPheValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLysA 536
 |||
 1617 TTTTGTGACGAGATCACTAAAGGCTCTACTGCTAAATTTGACGGTAAAG 1666
 536 lAlaArgGluIleLeuLysLysProValAlaLysMet 548
 |||
 1667 CAATTAGAGAAATATCTGAAGAAACGTTGCTPAAGATG 1704
 seq_name: /SID52/gcgdata/geneseq/geneseq/NA1996.DAT: AAT33850

seq_documentation_block:

ID AAT33850 standard; DNA: 2019 BP.

AC AAT33850;

DT 12-NOV-1996 (first entry)

DE DNA encoding streptavidin/luciferase fusion protein.

KW Streptavidin; luciferase; fusion protein;

KW Streptomyces avidinii; Luciola lateralis; firefly;

KW recombinant production; industry; ds.

OS Synthetic.

XX

XX Key

XX Location/Qualifiers

XX CDS

XX /*tag= a

XX JP07289264-A.

XX 07-NOV-1995.

XX 27-APR-1994; 94JP-0090275.

XX 27-APR-1994; 94JP-0090275.

XX (KIRK) KIRKMAN CORP.

XX WPI: 1996-015269/02.

XX P-PSDB; AAW04208.

XX New mutant streptavidin (SA) gene - fused to a firefly luciferase

XX gene, for the recombinant prepn. of a SA-FL fused protein

XX Example 1; Pages 5-6; 12pp; Japanese.

XX The present sequence encodes a streptavidin/luciferase fusion

XX protein, comprising the Streptomyces avidinii streptavidin gene

XX and the Luciola lateralis (firefly) luciferase gene. The fusion

XX protein can be prepd. by inserting the recombinant DNA encoding

XX it into a Escherichia species microbe, culturing the transformed

XX CC microbe in a medium and collecting the fusion protein from the

XX culture. The fusion protein can be used in industry.

SQ Sequence 2019 BP; 590 A; 412 C; 467 G; 550 T; 0 other;

alignment_scores:

Quality: 2799.00

Ratio: 5.126

Percent Similarity: 100.000

Percent Identity: 99.267

alignment_block:

US-09-581-241-6 x AAT33850

Align seg 1/1 to: AAT33850 from: 1 to: 2019

3 AsnMetGluAsnAspGluAsnIleValIyGlyProGluProPheTyrrPr 19

```

1282 CTCGATTAATATGATTTATCAAAATTTAGTTGAATTGCATCTGGCGGACC 1331
319 aProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheAsnLeuP 336
1332 ACCTTATCTAAAGAAATTTGGTGAAGCTGTGGTAGACGTTTAAATTAC 1381
336 roGlyValArgGlnGlyTyrGlyLeuThrGluThrSerAlaIleIle 352
1382 CGGGTGTCTGCAAGGCTATGTGTTTAAACAGAAACCACTTCGCAATTATT 1431
353 IleThrProGlnGlyAspAspLysProGlnAlaSerGlyLysValAlaPr 369
1432 ATCAACACGGAGCGCATGATTAACCAAGGCTGCTTGGCAAAATTTGGGCC 1481
369 oleuPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGlyP 386
1482 ATTAATTTAAAGCAAAAGTTATCGATCTTGATTAACAAAAAATTTGGGCC 1531
386 roAsnArgArgGlyGluValAlaLysGlyProMetLeuMetLysGly 402
1532 CGAACAGACGTGGAGAAAGTTTGTAAAGGCTCTTATGCTTATGAAGCT 1581
403 TyrValAspAsnProGlnAlaThrArgGluIleIleAspGlnGlyTyr 419
1582 TATGTAGATATATCCAGAGCAACAGAAATCATATGATGAAGAGTTTG 1631
419 pleuHisThrGlyAspIleGlyTyrTyrAspGlnGlyLysHisPhePheI 436
1632 GTTGCAACAGGAGATATGGTATTACATGAAGAAACAAATTTCTTTAA 1681
436 leValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGlnValPro 452
1682 TCGTGATGCTTGAAGCTCTTTAATCAATPACAAAGATATCAAGTACCA 1731
453 ProAlaGlnLeuGlnSerValLeuLeuGlnHisProAsnIlePheAspAl 469
1732 CCGCTGAATTAGAAATCTGTTCTTTTGCAACATCCAAATATTTTGATGC 1781
469 aGlyValAlaGlyValProAspProIleAlaGlyGlnLeuProGlyValAla 486
1782 CGGCGTGTGGCGTTCCAGATCCATACCTGCTGAGCTTCCGGGAGCTG 1831
486 aIleValAlaLeuLysLysGlyLysSerMetThrGluLysGluValMetAsp 502
1832 TTGTTGTAAGTCAAAAGGAAATCATGATCAAGAAAGAAAGTAAATGAT 1881
503 TyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyValAla 519
1882 TACGTTGCTAGTCAGATTTCAATGCAAAACGTTTGGCGTGGTGTCCG 1931
519 gpHeValAspGlnValProLysGlyLeuThrGlyLysIleAspGlyLysA 536
1932 TTTTGTGACGAAGTAACTAAAGGTCTCTACGTGTAATTTGACGGTAAAG 1981
536 laIleArgGlnIleLeuLysLysProValAlaLysMet 548
1982 CAATTAGAGAAATTAAGTGAAGAAACCAAGTTGCTAAGATG 2019

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT:AA133851
seq_documentation_block:
ID AA133851 standard; DNA: 2055 BP.
XX
AC AA133851;
XX
DT 12-NOV-1996 (first entry)
XX
DE DNA encoding mutant streptavidin/Luciferase fusion protein.
XX
KW Mutant; streptavidin; luciferase; fusion protein;
KW Streptomyces avidinii; Luciola lateralis; firefly;
KW recombinant production; industry; ds.
XX

```



```

419 pleuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisPheMet 436
      |||||||
1668 GTTGCACACAGGAGATATGGTATTCAGTATGAGAAAAACATTCTTTA 1717
436 ILeValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGlnValPro 452
      |||||||
1718 TCGTGATCGTTTGAAGCTCTTATCAATAAAGATATCAAGTACCA 1767
453 ProAlaGluLeuGluSerValLeuGlnHisProAsnIlePheAspAl 469
      |||||||
1768 CCTGCTGAATTCGAATCTGTTCTTTCACACATCAATAATTTTGTATGC 1817
469 aGlyValAlaGlyValProAspProIleAlaGlyGluLeuProGlyValAla 486
      |||||||
1818 CGGCGTTGCTGGGCTCCAGATCTTAGCTGTGAGCTCCGGAGCTG 1867
486 aValValLeuLysLysGlyLysSerMetThrGluLysGluValMetAsp 502
      |||||||
1868 TTTGTGTACTTGAAAAAGAAAAATCTATGACTGAAAAAGAAATATGAT 1917
503 TyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyValAla 519
      |||||||
1918 TACGTTGCTAGTCACTTCAATGCAAAACGTTGCGTGTGCTCCG 1967
519 gPheValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLysA 536
      |||||||
1968 TTTTGTGACGACGAACTAAAGCTCTCACTGCTAAATTTGACGGTAAAG 2017
536 ILeIleArgGluIleLeuLysLysProValAlaLysMet 548
      |||||||
2018 CAATTAGAGAAATACTGAAAGAACCGTGTCTAAGATG 2055
seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:AAV23580
seq_documentation_block:
ID AAV23580 standard; cDNA to mRNA; 2364 BP.
XX
AC AAV23580;
XX
DT 16-JUL-1998 (first entry)
XX
DE Antibody-Firefly Luciferase fusion protein gene.
XX
KW Firefly Luciferase; antibody-Luciferase fusion protein; ds.
XX
OS Luciola cruciata.
XX
FH Key 1. 2364
FT CDS
FT /tag= a
FT /transl_except= (pos: 670..672, aa: Glu)
FT /transl_except= (pos: 739..741, aa: Trp)
FT /transl_except= (pos: 1369..1371, aa: Ala)
FT /note= "no stop codon given"
XX
PN JP09187281-A.
XX
PD 22-JUL-1997.
XX
PF 09-JAN-1996; 96JP-0001812.
XX
PR 09-JAN-1996; 96JP-0001812.
XX
PA (KIKR ) KIKROMAN CORP.
XX
DR MPI; 1998-275089/25.
XX
DR P-PDB; AAW53882.
XX
PT Antibody-firefly Luciferase fused protein - and related products
PT i.e. firefly Luciferase fused gene, recombinant DNA and its
PT preparation
XX

```

```

PS Disclosure: Page 13; 17pp; Japanese.
XX
CC This sequence encodes the fusion protein of the invention. The protein is
CC a antibody-firefly luciferase fusion protein, in which an antibody part
CC consisting of a peptide having antibody activity is combined with an
CC enzyme part consisting of firefly luciferase.
XX
SQ Sequence 2364 BP; 697 A; 444 C; 559 G; 664 T; 0 other:

alignment_scores:
      Quality: 2798.00      Length: 545
      Ratio: 5.134      Gaps: 0
      Percent Similarity: 100.000      Percent Identity: 99.450

alignment_block:
US-09-581-241-6 x AAV23580 ..
Align seg 1/1 to: AAV23580 from: 1 to: 2364

4 MetGluAsnAspGluAsnIleValTyrGlyProGluProPheTyrProI 20
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
730 CTGCAAGACGATGAAAAATATGTGTATGTGCTCGAACCACTTTTACCCCTAT 779
20 eGluGluGlySerAlaGlyAlaGlnLeuArgLysTyrMetAspArgTyrA 37
      |||||||
780 TGAAGAGGATCTGCTGGAGCACAATGCCCANAGTATATGATGATGATG 829
37 ILeLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGlyValAspTyr 53
      |||||||
830 CAAACTGTGAGACGAAATGCTTTACTAACGCACTTACCGGTGCGATAT 879
54 ThrTyrAlaGluTyrLeuGluLysSerCysLeuGlyGluAlaLeuLys 70
      |||||||
880 ACCTACGCCGAATCTTACGAAAATCATCTGCTGACGAGAGGCTTTAA 929
70 sAsnTyrGlyLeuValAlaAspGlyValArgIleAlaLeuCysSerGluAsn 87
      |||||||
930 GAATATATGTTGTGTTGTGTATGGAAGAAATGCGTTATGACATGCAAAACT 979
87 yGluGluPhePheIleProValLeuAlaGlyLeuPheIleGlyValGly 103
      |||||||
980 GTGAAGAGTTCTTATTCCTGATTAAGCCGTTTATTTATAGTGTGCGCT 1029
104 ValAlaProThrAsnGluIleTyrThrIleuArgGluLeuValHisSerLe 120
      |||||||
1030 GTGGCTCCCACTAATGAGATTTACACTCTACGTGAATTTGTTACACAGTTT 1079
120 uGlyIleSerLysProThrIleValPheSerSerLysGlyLeuAspL 137
      |||||||
1080 AGGCATCTCTTAAGCCACAAATGTCTTATGTTCTTAAAGATTAAGATA 1129
1080 AGGCATCTCTTAAGCCACAAATGTCTTATGTTCTTAAAGATTAAGATA 1129
137 yValIlePheThrValGlnLysThrValAlaIleLysThrIleValIle 153
      |||||||
1130 AAGTTATATACGTACAAAAAACGTAACGTCTTAAACCTTTGTTANA 1179
154 LeuAspSerLysValAspTyrArgLysTyrGlnSerMetAspAsnPheI 170
      |||||||
1180 TTGCACAGCAAAAGTGAATTAATGAGTATATCAATCAATGACAACTTTAT 1229
170 eLysLysAsnThrProGlnGlyPheLysGlySerSerPheLysThrValG 187
      |||||||
1230 TAAAAAAAACACTCCACAAAGTTTCAAAAGATCAAGTTTAAACCTTAG 1279
187 IuValAsnArgLysGluGlnValAlaLeuIleMetAsnSerSerLysSer 203
      |||||||
1280 AAGTTAACCGCAAAAGACAAAGTTGCTCTTATATGAACTCTTCGGGCTCA 1329
204 ThrGlyLeuProLysGlyValAlaGlnLeuThrHisGluAsnIleValThrA 220
      |||||||
1330 ACCGGTTTGCCAAAAGTGTGCAACTTACTCATGAATAATTTGTTACGCG 1379
220 gPheSerHisAlaArgAspProIleTyrGlyAsnGlnValSerProGlyT 237

```

```

1380 TTTTCTCAGCTAGAGATCCATTATGAAACCAAGTTTCAACGCGCA 1429
237 hrAlaIleuThrValValProPheHisIsglyPheGlymetPheThr 253
1430 CGGCTATTCTAGTACTAGTACCATTCATCATGTTTGGTATGTTTACT 1479
254 ThrLeuGlyTyrIleuThrCysGlyPheArgIleValMetLeuThrLysPh 270
1480 ACTTTAGGCTATCTACTGTTGGTTTTCGATATGTCATGTTAAACGAAT 1529
270 eaSPGluGluThrPheLeuLysThrLeuGlnAspTyrLysCysSerSerV 287
1530 TGAACGAGAGACTTTTAAACACGTCGACAGATTCAAAATGTTCAAGCG 1579
287 alIleuValProThrIleuPheAlaIleuAsnArgSerGluLeu 303
1580 TTTATCTTGTGACGACTTGTGTCATTCCTTAATAGAGTGAATTAATC 1629
304 AspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGlyLysAlaPr 320
1630 GATAAATATATTTATCAATTTAGTTGAATTCATCTGGCGACGACAC 1679
320 olSerLysGluIleGlyLysAlaValAlaArgArgPheAsnLeuProG 337
1680 TTTATCTTAAAGAAATGTCGACACTGTGCTACACGTTTAAATTTACCG 1729
337 LysValArgGlyTyrGlyLeuThrGluThrSerAlaIleIleIle 353
1730 GTGTTGCTCAAGGCTATGTTTAAACAGAAACACTCTGCAATTAATATC 1779
354 ThrProGluLysAspLysProGlyAlaSerGlyLysValAlaProle 370
1780 ACACCGGAGGCGGTGATTAACCAAGCTCTCTGGCAAACTGGCCATT 1829
370 upheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGlyPro 387
1830 ATTTAAAGCAAAAGTATCGATCTTGATCTAAATAAACTTGGCCCGA 1879
387 snArgArgGlyLysValLysValLysGlyProMetLeuMetLysGlyTyr 403
1880 ACACACGTGGAGAGTTGTGTTAAAGGTCCTATGCTTAAGAAAGGTTAT 1929
404 ValAspAsnProGluAlaThrArgGluIleIleAspGluGluTyrPle 420
1930 GTAAATATCCAGAACACACAGGAAATCATCATGAAGAGTTGTT 1979
420 uHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisPhePheIleV 437
1980 GCACACAGGAGATATGAGTATCATGATGAAGAAACATTTCTTATCG 2029
437 alaAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGluValProPro 453
2030 TGGATCGTTTAAAGCTTTAATCAAAATACAAAGATATCAAGTCCACCT 2079
454 AlaGluLeuGluSerValIleuLeuGlnHisProAsnIlePheAspAlaG 470
2080 GCTGATTAAGATCTGTTCTTTGCAACATCCAAATATTTTGAATGCCG 2129
470 yValAlaGlyValProAspProIleAlaGlyLysLeuProGlyAlaVal 487
2130 CGTTGCTGGCTGCCAGATCTATAGCTGGAGCTTCCGGAGACTGTG 2179
487 alValLeuLysLysGlyLysSerMetThrGluLysGluValMetAspTyr 503
2180 TTGTACTTGAAGAAAGAAATCTATGACTGAAAAGAGTATGATGATTAC 2229
504 ValAlaSerGluValSerAsnAlaLysArgLeuArgGlyLysValArgPh 520
2230 GTTGCTAGTCAAGTTTCAATGCAAAAGCTTGGCGTGTGTCCTGTT 2279
520 eValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLysAlaI 537

```

```

2280 TGTGACGACGAGTACCTAAAGGCTCCTACTGTAATAATTGACGCTAAAGCA 2329
537 leArgGluIleLeuLysLysProValAlaLysMet 548
2330 TTAGAGAAATCTGAAAGAACCAAGTTGCTAAGATG 2364
seq_name: /SDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AA25717
seq_documentation_block:
ID AAX25717 standard; cDNA to mRNA; 1656 BP.
XX
AC AAX25717;
XX
DF 21-MAY-1999 (first entry)
XX
DE Firefly luciferase gene #3.
XX
KW Bioluminescent protein; catalytic efficiency; stability; firefly;
KW luciferase; chimeric; Luciola cruciata; Luciola lateralis;
KW Photinus pyralis; primer; PCR; amplification; ss.
XX
OS Chimeric - Luciola lateralis.
XX
OS Chimeric - Photinus pyralis.
XX
W09902697-A1.
XX
PD 21-JAN-1999.
XX
PF 30-JUN-1998; 98WO-JP02936.
XX
PR 08-JUL-1997; 97US-0051917.
XX
PA (KIKK) KIKKOMAN CORP.
XX
PI Hirokawa K, Kajiyama N, Murakami S;
XX
DR WPI: 1999-120898/10.
XX
DR P-PSDB: AAW993366.
XX
PT New bioluminescent protein with improved properties - has greater
PT catalytic efficiency and stability and is obtained by modification
PT of natural precursors
XX
PS Example 3; Page 30-31; 53pp; Japanese.
XX
CC The invention relates to the generation of bioluminescent proteins with
CC improved catalytic efficiency and stability. The proteins are generated
CC by addition, deletion or substitution of amino acids from a precursor
CC protein e.g. firefly luciferases, or by constructing chimeric luciferase
CC proteins from luciferase genes from e.g. Luciola cruciata, L. lateralis
CC and Photinus pyralis. This sequence represents an example of a chimeric
CC luciferase gene of the invention.
XX
SQ Sequence 1656 BP; 529 A; 283 C; 366 G; 478 T; 0 other;
XX
alignment_scores:
Quality: 2701.00 Length: 543
Ratio: 5.039 Gaps: 0
Percent Similarity: 98.711 Percent Identity: 96.133
alignment_block:
US-09-581-241-6 x AAX25717 ..
Align seg 1/1 to: AAX25717 from: 1 to: 1656
1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17
1 ATGGAAGAAATGAGAACGATGAAATATTTGTATGCTGCTTAACCAAT 50
17 eTyrProIleGluGluGlySerAlaGlyAlaGluLeuArgLysTyrMet 34
51 TTACCTATTGAAAGAGGATCTGCTGAGACACAAATGCGCAATATATG 100

```

34 sPaRgTYrAlAlaLysLeuGlyAlAlaLeAlaPheThrAsnAlaLeuThrGly 50
 101 ATCGATATGCAAAACTTGGAGCAATGCTTTTACTAGCAAGCACTTACCGGT 150
 51 ValAspTYrThrTYrAlaGluTYrLeuGluLysSerCysCysLeuGly 67
 151 GTGATTTATAGCTAGCGCCGAAATCTTAGAAAAATCATGCTGCTAGGACA 200
 67 uAlaLeuLysAsnTYrGlyLeuValValAspGlyArgIleAlaLeuCys 84
 201 GGCTTTAAAGAAATTAGTGTGCTTCTTATGAGAAATTCGCTTAGCA 250
 84 eRGLAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPhe 100
 251 GTGAAACCTGAGAAATCTTATCTCCGTATAGCCGTTATTATTATA 300
 101 GlyValGlyValAlaProThrAsnGluIleTYrThrLeuArgGluLeuVal 117
 301 GGTCGCGGTGGCTCCAACTAATGAGATTACACTACGTAGCAATTTGGT 350
 117 IHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysG 134
 351 TCACAGTTTAGGCATCTCTAACCCAAATGTATTTACTTCTTAAAG 400
 134 LysLeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
 401 GATTAGATAAAGTTATTAAGTACCAAAAAACCGTAACTGCTATTAAACC 450
 151 IleValIleLeuAspSerLysValAspTYrArgGlyTYrGlnSerMetAs 167
 451 ATTGTTATATGACAGCAAGTGCATTTATGAGTTATCAATCAATGGA 500
 167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerPheL 184
 501 CAACCTTTATTAATAAAACACCTCCACAAGGTTTCAAGAGATCAAGTTTA 550
 184 YSThrValGluValAsnArgLysGluGlnValAlaLeuIleMetLysSer 200
 551 AAACGTGAGAGTTACCGCAAGAACAAAGTTGCTTTATTAATGAACTCT 600
 201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnI 217
 601 TCGGGTTCACACCGGTTTGCAGAAAGGTGCAACTTACTCATGAAATTT 650
 217 eValThrArgPheSerHisAlaArgAspProIleTYrGlyAsnGlnValS 234
 651 GGTCAGTAGATTTTCTCAGCTAGAGATCCAATTATGTGAAACCAAGTTT 700
 234 eRProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
 701 CACGAGGACGGCTATTTTAACGTAGTACCATTCATCATGATGGTTTGGT 750
 251 MetPheThrThrLeuGlyTYrLeuThrCysGlyPheArgIleValMetLe 267
 751 ATGTTTACTACTTAGGCTATCTAACTGTGTTTCTGATTTGCATATTT 800
 267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTYrLysC 284
 801 AAGCAAAATTTGACAGAGAGACTTTTAAAAACACGTCAAGATTACAAAT 850
 284 YSSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
 851 GTTCAACGCTTATCTTGACCGACTTTGTTGCAATTTCTTAATAGAACT 900
 301 GlnLeuLeuAspLysTYrAspLeuSerAsnLeuValGluIleAlaSerGly 317
 901 GAATTTACTCATTAATATGATTTATCAAAATTTAGTGAATTTGATCGTGG 950
 317 YGLValAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheA 334
 951 CGAGACACCTTTATCTAAGAAATTTGGTGAAGCTGTGCTTAGACGTTTAA 1000

334 snLeuProGlyValAlaArgGlnGlyTYrGlyLeuThrGluThrThrSerAla 350
 1001 ATTTACCGGGGTGTTCTCAAGGCTATGTTTAAACAGAAACAACTCTGCA 1050
 351 IleIleIleThrProGluLysAspLysProGlyAlaSerGlyLysVal 367
 1051 ATTATATCAACCGGAGCGCATGATTAACAGCGTCTTGGCAAAAT 1100
 367 lValProLeuPheLysAlaLysValIleAsnLeuAspThrLysLysThrL 384
 1101 TGTGCCATTTATTTAAAGCAAAAGTTATTCATCTGTGATCTTAAAAAACT 1150
 384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
 1151 TGGGCCCGCAACAGACCTGAGAAAGTTGTGTAAGGGTCTTATGCTTAG 1200
 401 LysGlyTYrValAspAsnProGluAlaThrArgGluIleIleAspGlu 417
 1201 AAAGTTATGTAGATTAATCCAGAACGACAAAGAAATCATGATGACAGA 1250
 417 uGlyTYrLeuHisThrGlyAspIleGlyTYrTYrAspGluGluLysHisP 434
 1251 AGTTGTTTGACACAGAGATATGGGTATTACATGAGTGAAGAAAAACATT 1300
 434 hePheIleValAspArgLeuLysSerLeuIleLysTYrLysGlyTYrGln 450
 1301 TCTTATATGTCGATCGTTTGAAGTCTTTATCAATACAAAGATATCAG 1350
 451 ValProProAlaGluLysGluSerValLeuLeuGlnHisProAsnIlePh 467
 1351 GTGGCCCCCGCTGAATTTGAAATGCAATATTGTTTCAACACCCCAATCT 1400
 467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
 1401 CGACGCGGCGCTGCGCAGGCTTCCGACGATGACCCCGGTGAACTTCCG 1450
 484 lYAlaValValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
 1451 CCGCCGTTGTTGTTTGGAGCACGGAAGACGATGACGGAAGAAACAGATC 1500
 501 MetAspTYrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
 1501 GTGGATTACGTCGCAAGTCAAGTAAACACCGGAAAGTTGCGGAGAG 1550
 517 YValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
 1551 AGTTGTTTGTGACGAAAGTAAAGGTCCTTACGCGAAAACTCGACG 1600
 534 lYlYAlaIleArgGluLeuLysLys 543
 1601 CAGGAAATCAAGAGATCTCTCATTAAG 1629
 seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT.AAN91170
 seq_documentation_block:
 ID AAN91170 standard; DNA; 1644 BP.
 XX
 AC AAN91170;
 XX
 DT 14-JUN-1990 (first entry)
 XX
 DE Recombinant luciferase gene.
 XX
 KW Luciferase; ATP assay; ss.
 XX
 OS Luciola cruciata.
 XX
 PN EP301541-A.
 XX
 XX 01-FEB-1989.
 PD
 XX
 PF 28-JUL-1988; 88BP-0112233.
 XX

PR 29-JUL-1987; 87JP-0187724.
PR 29-JUL-1987; 87JP-0187725.
PR 20-AUG-1987; 87JP-0205194.

XX
XX (KIKK) KIKKOMAN CORP.

PI Masuda T, Tatsumi H, Nakano E;

DX WPI; 1989-03443/05.

DR P-PSDB; AAP94367.

XX
XX Recombinant luciferase gene - used for efficiently producing luciferase
PT in E. coli, obtaining a stable and highly active prod.

XX
XX Claim 2; fig 3; 35pp; English.

CC The sequence was obt. from cDNA prep. from mRNA isolated from
CC L. cruciata tails. The DNA can be incorporated into vectors for
CC transformation of E. coli JM 1010.

XX
XX Sequence 1644 BP; 529 A; 276 C; 338 G; 501 T; 0 other;

alignment_scores:

Quality: 2684.00 Length: 548
Ratio: 4.943 Gaps: 0
Percent Similarity: 99.088 Percent Identity: 93.431

alignment_block:

US-09-581-241-6 x AAN91170 ..

Align seg 1/1 to: AAN91170 from: 1 to: 1644

1 MetGluAsnMetGluAsnAspGluAsnIleValIYrGlyProGluProph 17
1 ATGGAACACATGGAACACATGGAACATATGTAGTTGACCTAAACCGT 50
17 eYrProIleGluGluGlySerAlaGlyAlaGluLeuArgIYrMetA 34
51 TTACCTATCGAAGAGATGCTGTGGAACACAAATACGCAAAATACATGG 100
34 sPArGTYrAlaIYsLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 AGCGATATGCAAAACTTGGCGCAATTTGCTTTACAAATGCAAGTACTGCT 150
51 ValAspTYrThrTYrAlaGluTYrLeuGlySerCysGlyLeuGlyG 67
151 GTTGATTAATCTTACCGCAATATGGAATATCATGTTGCTTAGGAAA 200
67 uAlaLeuAsnTYrGlyLeuValValAspGlyrAlaIleAlaLeuGly 84
201 ACCTTTGCAAAATATGTTGGTTGTGTGATGGCAAAATGCGTTATGCA 250
84 eRGLuAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
251 GTGAAACACGTGAAGATTTTATTCCTGTAATAGCCGACTGTTATA 300
101 GlyValGlyValAlaProThrAsnGluIleTYrThrLeuArgGluLeu 117
301 GGTTGTGCTTGCAACCCACTAATGAGATTACACTTACGTAACGTAACG 350
117 IHSerLeuGlyIleSerIYsProThrIleValPheSerIYsLysG 134
351 TCAAGTTAGGTATCTTAACACACAACTTATTTAGTTCTTAAAAAAG 400
134 IYLeuAspIYsValIleThrValGlnIYsThrValThrAlaIleIYsThr 150
401 GCTTAAATAAGTTATTAACAGTAAACAAACAGTAACTACTATTAAC 450
151 IleValIleLeuAspSerIYsValAspTYrArgGlyTYrGlnSerMet 167
451 ATTGTATCTAGTATGCAAAAGTTGATTATGAGGATATCAATGTCCTGA 500

167 pAsnPheIleIYsAsnThrProGlnIYpHeIYsGlySerSerPheI 184
501 CACCTTATTAAGAAAGACACCTCACAGGTTTTCAGCTCAGTTTCA 550
184 YsThrValGluValAsnArgIYsGluGlnValAlaLeuIleMetAsnSer 200
551 AAACGTGGAGAGTTGACCGTAAAGAACAGTTGCTCTTATTAATGAACCT 600
201 SerGlySerThrGlyLeuProIYsGlyValGlnLeuThrHisGluAsnI 217
601 TCGGTTCTTACCGGTTGCCAAAAGCGTACAACTTACACAGAAATATC 650
217 eValThrArgPheSerHisAlaArgAspProIleTYrGlyAsnGlnVal 234
651 AGTCACTAGATTTTCTATGCTAGATCGATTCGATTCGTAACCAAGTT 700
234 eRProGlyThrAlaIleLeuThrValValProPheHisIleGlyPheGly 250
701 CACAGGACACCGCTGTTTAACTGCTGTTCCATTCATCATGATGTTGGT 750
251 MetPheThrThrLeuGlyTYrLeuThrCysGlyPheArgIleValMetI 267
751 ATGTTCACTACTCTAGGATTAATTTGTTGTTGTTGTTGTTGTTGTT 800
267 uThrIYsPheAspGluGluThrPheLeuIYsThrLeuGlnAspTYrLysC 284
801 AACCAAAATTCATGAGAACAACTTTTAAACCTCTACAAAGATTATAAT 850
284 YsSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTACAACTGTTATCTTGTACCACTGTTGTTGCAATTCACAAACAAAGT 900
301 GluLeuAsnAspIYsTYrAspIYsSerAsnLeuValGluIleAlaSerG 317
901 GAATTTACTCAATAATTCAGATTGTCTCAATTTAGTTGAGTTCACTGG 950
317 YGlyAlaProLeuSerIYsGluIleGlyGluAlaValAlaArgArgPhe 334
951 CGGAGCACTTTATCAAAAGAGTGTGTAAGCTGTTGCTAGACGCTTTA 1000
334 snLeuProGlyValArgGlnIYrGlyLeuThrGluThrThrSerAla 350
1001 ATCTTCCCGGTTGCTGCAAGGTTATGTTTAAACAGAAACATCTGCC 1050
351 IleIleIleThrProGluGlyAspAspIYsProGlyAlaSerGlyIYs 367
1051 ATTATTTATCACCCAGAGAGACGATTAACCAAGACCTCTGGAACACT 1100
367 IValProLeuPheIYsAlaIYsValIleAspLeuAspThrIYsLysThr 384
1101 CGTGCCGTTGTTAAAGCAAAAGTTATTTGATCTGATACCAAAAAATCTT 1150
384 eUGlyProAsnArgArgGlyGluValIYsValIYsGlyProMetLeuMet 400
1151 TAGTCTCCAAACAGAGTGAGAAAGTTGTGTTAAAGAACCTTAAGCTTATG 1200
401 LysGlyTYrValAspAsnProGluAlaIleThrArgGluIleIleAspGlu 417
1201 AAAGTTATGTAATTAATCCGAAGCAACAAAGAACTTATGACGAAGA 1250
417 uGlyTrpLeuHisThrGlyAspIleGlyTYrTYrAspGluGluIYsHisP 434
1251 AGTTGCGCTGCACACCGAGATATTGATTTATGATGAAGAAAAACATT 1300
434 hePheIleValAspArgLeuIYsSerLeuIleIYsTYrIYsGlyTYrGln 450
1301 TCTTATTTGTCATGCTGTTGAAGCTTTTAAATCAATCAAAAGGTTATAC 1350
451 ValProAlaGluLeuGluIYsValLeuLeuGlnIHisProAsnIlePh 467
1351 GTACCACTGCGCAATTTGAATCCGTTCTTTTGCACATCATCATCTT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484

```

|||||
1401 TGAATGCTGGTGTCCGCGCTCTGATCCTGTAGCTGCGAGCTTCAG 1450
484 1yAlaValValleuLysGlyLysSerMetThrGluLysGluVal 500
1451 GAGCCGTTGTTGTTACTGGAAGGAAAAATATGACCGAAAAAGAGTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgLysI 517
1501 ATGATATTATGTTCAAGTCAAGTTTCAATGCAAAACGTTTACGTGCTG 1550
517 yAlaAlaGpHeValAspGluValProLysGlyLeuThrGlyLysIleAsp 534
1551 TGTTCGTTTGTGATGAAGTACCTAAAGCTTACTGGAATAATTGACG 1600
534 1yLysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
1601 GCAGACCAATTAGAGAAATCCTTAAGAACCACTGCTTAGATG 1644

```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:AAQ03801

seq_documentation_block:

```

ID AAQ03801 standard; cDNA; 1644 BP.
XX
AC AAQ03801;
XX
DT 22-AUG-1990 (first entry)
DT 26-FEB-1993 (revised entry)
XX
DE Recombinant luciferase gene.
XX
KW Luciferase; ATP; E. coli; photon; ds.
OS Synthetic.
XX
PN JP02065780-A.
XX
PD 06-MAR-1990.
XX
PF 01-SEP-1988; 88JP-0216229.
XX
PR 01-SEP-1988; 88JP-0216229.
XX
PA (KIKK ) KIKKOMAN CORP.
XX
DR WPI; 1990-113360/15.
DR P-PSDB; AAR05788.
XX
PT Prepn. of luciferase -
PT using Escherichia sp. bacteria contg. recombinant DNA.
PS Claim 2; Page 416 + Fig 3; 20pp; Japanese.
XX
XX Prepn. of luciferase comprises culturing Escherichia sp. contg.
XX recombinant DNA, and harvesting luciferase from the culture medium.
XX The recombinant DNA is a vector DNA into which has been inserted the
XX luciferase gene which has the nucleotide sequence below. Luciferase
XX can be produced quickly and efficiently with an increased photon
XX output, useful in assaying ATP.
XX
S0 Sequence 1644 BP; 529 A; 276 C; 338 G; 501 T; 0 other;

```

alignment_scores:

```

Quality: 2684.00      Length: 548
Ratio: 4.943          Gaps: 0
Percent Similarity: 99.088      Percent Identity: 93.431

```

alignment_block:

us-09-581-241-6 x AAQ03801 ..

Align seg 1/1 to: AAQ03801 from: 1 to: 1644

```

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProph 17
|||||
1 ATGGAACAACTGGAAAAAGATGAAATATTTGATGTTGACCTTAACCGTT 50
17 eTyrProIleGluGluGlySerAlaGlyIleValArgLysTyrMet 34
|||||
51 TTACCCCTATCGAAGAGGATCTGCTGGACACAACTTACGCAATACATGG 100
34 sParGlyTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
|||||
101 AGCATTATGCAAAACTTGGCGCAATTGCTTTTACAAATGACGTTACTGCT 150
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysGlyLeuGly 67
|||||
151 GTTGATTATTCTTACCGCCGAAATCTTGAGAAATCATGTTGCTTAGGAAA 200
67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyTyrGlyIleAlaLeuCys 84
|||||
201 ACCTTGCAGAAATTTATGTTGTTGTTGATGGCAGAAATTCGTTATGCA 250
84 erGluAsnGlyGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
|||||
251 GTGAAACCTGTGAAGATTTTATTCCTGTATATACCGGACGTTTATA 300
301 GGTGTAGGTGTTGCACCACTATGATGATTTACCTTTACGTAACACTGG 350
117 IHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysGly 134
|||||
351 TCACAGTTTAGATATCTTAACCAACAATTTGATTTAGTTCTTAAAAAAG 400
134 1yLeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
|||||
401 GCTTACATAAAGTTATTAACAGTACAGAAAAACAGTACTATTTAAAC 450
151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAs 167
|||||
451 ATTGTTATCTAGATAGCAAGTTGATTCGAGATATCAATCAATGCTGGA 500
167 pAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPhe 184
|||||
501 CACCTTATTAATAAGAAACACCTCCACAGGTTTCAAGATCCAGTTTCA 550
184 ySThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
|||||
551 AAACGTGGAAGTTGACCGTAAAGAACAGAGTTGCTTTATTAAGACTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnI 217
|||||
601 TCGGGTTCTACCGGTTTGCCAAAAGCGTACAACTTACACGAAAAATAC 650
217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
|||||
651 AATCCTACATATTTCTCATGCTAGAGATCCGATTTATGGTAAACAGTTT 700
234 exProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
|||||
701 CACAGAGCACCGCTGTTTAACTGCTGCATTCATCATCATGCTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
|||||
751 ATGTTCACTACTCTAGGGTATTTTAAATTTGCTGTTTCTGTTGTAATGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
|||||
801 AACCAAAATTCGATGAAGAAACATTTTAAAAACCTTACAGATTTTAAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
|||||
851 GTACAGAGTGTATTTCTGTACCGACCTGTTTGAATTCCTCAACAAAGT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317

```

```
|||||.....  
901 GAATTACTCAATAAATGCAATTGTCAATTTAGTTCATCTCG 950  
317 YGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334  
|||||.....  
951 CGGAGCACCTTTATCAAGAAGAGTGGTAGAGCTTGTAGACGCTTAA 1000  
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrSerAla 350  
|||||.....  
1001 ATCTTCCGGTTCGTCGAAGGTTATGTTTAACAGAAACACATCTGCC 1050  
351 IleIleIleThrProGluGlyAspAspLysProGlyValAspGlyLysVa 367  
|||||.....  
1051 ATTATTTATTCACCGAAGAGACAGATAAACAGAGACTTCTGAAAGT 1100  
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384  
|||||.....  
1101 CGTCCCGTTGTTAAAGCAAAAGTATTGATCTGATACCAAAATCTT 1150  
384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400  
|||||.....  
1151 TAGGTCCCTAACAGACGTGAGAAAGTTGTGTTAAAGAGACCTATGCTATG 1200  
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGluG 417  
|||||.....  
1201 AAAGGTTATGTAAATTAATCCAGAAAGCAACAAAGAACTTATGACGAAGA 1250  
417 uGlyTyrPheuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434  
|||||.....  
1251 AGGTGGCTGCACACCGGATATTTGATATTGATGAGAAACAAACATT 1300  
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450  
|||||.....  
1301 TCTTTATGTGATCGTTGAAGCTTTAATCAATACAAAGGATACCAA 1350  
451 ValProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467  
|||||.....  
1351 GTACCACTGCGCAATTTGATCCGTTCTTTGCACATCCATCTATCTT 1400  
467 eaSpAlaGlyValAlaGlyValProAspProIleAlaGlyLeuLeuProG 484  
|||||.....  
1401 TGATGCTGTGTGGCGGCTTCCGTGATCCTGAGCTGGCGAGCTTCAG 1450  
484 lValAlaValValLeuLysLysGlyLysSerMetThrGluLysGluVal 500  
|||||.....  
1451 GAGCCGTTGTGTACTGGAAGCGGAAAAAATATGACCGAAAAAGAGTA 1500  
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyG 517  
|||||.....  
1501 ATGATTTATGTTGCAAGTCAGATTCAATGCAAAACGTTTACGTGTGG 1550  
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534  
|||||.....  
1551 TCTTGCTTTTGGATGAAGTACCTAAAGGTCCTTACTGAAAAAATTGAGC 1600  
534 lYlyAlaIleArgGluIleLeuLysLysProValAlaLysMet 548  
|||||.....  
1601 GCAGAGCAATTAGAGAAATCCTTAAGAAACGTTGCTTAGATG 1644
```


CC This sequence is a Luciola lateralis luciferase of the invention.
 CC The protein is a surfactant-tolerant luciferase which is derived from
 CC firefly luciferase by replacement of the glutamic acid residue at
 CC position 490 by another amino acid. The luciferase can be used in an
 CC assay of intracellular ATP in biological samples containing cells, for
 CC example for establishing the presence of cells, and their quantity, in
 CC agrochemicals, beverages and environmental samples. The activity of the
 CC surfactant-tolerant luciferase is not lowered by the presence of high
 CC concentrations of surfactant and therefore an assay of high accuracy can
 CC be carried out in a simple manner.

XX
 XX
 SQ Sequence 548 AA;

Query Match 100.0%; Score 2823; DB 20; Length 548;
 Best Local Similarity 100.0%; Pred. No. 1.1e-235;
 Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MENNENDENTVYGPPEPIPIEGSAGAOIKRYMDRYAKGAIATNALTGVDYTYAEYLE 60
 DB 1 memmendenlvyspepiyleegsagaqlrkymdryaklgatlnatlgvdytyaeyle 60
 QY 61 KSCCLGALKNYGLVVDGRITALCSENCEEFIPVLGLFVGAVAPRNEITLTRELVHSL 120
 DB 61 kscclgalknnyglvvdgritalcsenceeffipvlgflfvgavaprnelyltrelvhs1 120
 QY 121 GISKPTIVSSKKGLDKVITVOKTVAIKTIVILDSKVDYRGYQSDNFIKNTPOGFGK 180
 DB 121 giskptivsskkgldkvitvoktvaiktivildskvdyrgyqsdnfiikntpgfgk 180
 QY 181 SSFKTEVNRKEOVALIMNSSGSTGLPKGYOLTHEMLVTFESNARDITYGNVSPGTAIL 240
 DB 181 ssfktevnrkeovalimnssgstglpkgyolthelmlvtfeshardipygnvsgptail 240
 QY 241 TVVPFHGFGMETTLGYLTGCFRIWMLTKFDEETFLKTLQDYKSSVILVPTLFAILNRS 300
 DB 241 tvvpfhgfgmettlgyltcgfrimlwkfdeetflktdykcvsilvptlfaalnrs 300
 QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRNLPVGRGCGLTETTSAILITPESGDK 360
 DB 301 elldkydlsnlveiasgapskeigeavarrihlpvrgyglteftsaailitpegddk 360
 QY 361 PGASGVVPLFRKAVIDLDTKKTLPNRRGEVGVKGMKMGVVDNEARETIDEEGWL 420
 DB 361 pgasgvvplfrkavidldtkktlpnrrgevcvkgpmlmkgyvdnearetideegwl 420
 QY 421 HNGDICYDEEKHFETVDRKSLIKYKGYVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
 DB 421 hngdicydeekhfetvdrkslikykgyvppaellesvllqhpni fdagvagvpdiag 480
 QY 481 ELPGAVVVLLKKKSMTEKEVMDYVASOVSNAKRLRGVRFVDEVPKGLTGKIDGKAIREI 540
 DB 481 elpgavvvllkkksmtekevmdyvasvsnakrlrgvrfvdevpkgltgkldgkaiirei 540
 QY 541 LKKPVAKM 548
 DB 541 lkkpvakm 548

RESULT 2

AAAY22183 ID AAY22183 standard; Protein; 548 AA.

XX AC AAY22183;

XX DT 09-SEP-1999 (first entry)

XX DE L. lateralis luciferase protein sequence.

XX KW Luciferase; surfactant-tolerant; firefly; intracellular ATP assay.

XX OS, Luciola lateralis.

XX
 PN W09933997-A1.
 PD 08-JUL-1999.
 XX
 PF 24-DEC-1998; 98WO-JP05864.
 PR 26-DEC-1997; 97JP-0361022.
 XX
 PA (KITK) KIKKOMAN CORP.
 XX
 PI Hattori N, Murakami S;
 XX
 DR WPI: 1999-419109/35.
 DR N-PSDB: AAX84396.
 XX
 PT Modified luciferase tolerant to surfactants and useful for assay of
 PT intracellular ATP
 XX
 PS Claim 5; Page 43-48; 56pp; English.

CC This sequence is a Luciola lateralis luciferase of the invention.
 CC The protein is a surfactant-tolerant luciferase which is derived from
 CC firefly luciferase by replacement of the glutamic acid residue at
 CC position 490 by another amino acid. The luciferase can be used in an
 CC assay of intracellular ATP in biological samples containing cells, for
 CC example for establishing the presence of cells, and their quantity, in
 CC agrochemicals, beverages and environmental samples. The activity of the
 CC surfactant-tolerant luciferase is not lowered by the presence of high
 CC concentrations of surfactant and therefore an assay of high accuracy can
 CC be carried out in a simple manner.

XX
 XX
 SQ Sequence 548 AA;

Query Match 99.9%; Score 2821; DB 20; Length 548;
 Best Local Similarity 99.8%; Pred. No. 1.7e-235;
 Matches 547; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MENNENDENTVYGPPEPIPIEGSAGAOIKRYMDRYAKGAIATNALTGVDYTYAEYLE 60
 DB 1 memmendenlvyspepiyleegsagaqlrkymdryaklgatlnatlgvdytyaeyle 60
 QY 61 KSCCLGALKNYGLVVDGRITALCSENCEEFIPVLGLFVGAVAPRNEITLTRELVHSL 120
 DB 61 kscclgalknnyglvvdgritalcsenceeffipvlgflfvgavaprnelyltrelvhs1 120
 QY 121 GISKPTIVSSKKGLDKVITVOKTVAIKTIVILDSKVDYRGYQSDNFIKNTPOGFGK 180
 DB 121 giskptivsskkgldkvitvoktvaiktivildskvdyrgyqsdnfiikntpgfgk 180
 QY 181 SSFKTEVNRKEOVALIMNSSGSTGLPKGYOLTHEMLVTFESNARDITYGNVSPGTAIL 240
 DB 181 ssfktevnrkeovalimnssgstglpkgyolthelmlvtfeshardipygnvsgptail 240
 QY 241 TVVPFHGFGMETTLGYLTGCFRIWMLTKFDEETFLKTLQDYKSSVILVPTLFAILNRS 300
 DB 241 tvvpfhgfgmettlgyltcgfrimlwkfdeetflktdykcvsilvptlfaalnrs 300
 QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRNLPVGRGCGLTETTSAILITPESGDK 360
 DB 301 elldkydlsnlveiasgapskeigeavarrihlpvrgyglteftsaailitpegddk 360
 QY 361 PGASGVVPLFRKAVIDLDTKKTLPNRRGEVGVKGMKMGVVDNEARETIDEEGWL 420
 DB 361 pgasgvvplfrkavidldtkktlpnrrgevcvkgpmlmkgyvdnearetideegwl 420
 QY 421 HNGDICYDEEKHFETVDRKSLIKYKGYVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
 DB 421 hngdicydeekhfetvdrkslikykgyvppaellesvllqhpni fdagvagvpdiag 480
 QY 481 ELPGAVVVLLKKKSMTEKEVMDYVASOVSNAKRLRGVRFVDEVPKGLTGKIDGKAIREI 540

Db 481 elpavvvllkkgksmtekevmdyvaasvsnakrlrgvrtivdevpkqltqkldgkaltrei 540
 Oy 541 LKKPVAKM 548
 Db 541 lkkpvakm 548

RESULT 3

AAW04212 ID AAW04212 standard; Protein: 548 AA.

AC AAW04212;

DT 12-NOV-1996 (first entry)

DE Luciola lateralis luciferase.

KM Mutant; streptavidin; luciferase; fusion protein;

KM Streptomyces avidinii; Luciola lateralis; firefly;

OS recombinant production; Industry.

PN Luciola lateralis.

PD JP07289264-A.

XX 07-NOV-1995.

PF 27-APR-1994; 94JP-0090275.

PR 27-APR-1994; 94JP-0090275.

PA (KIKK) KIKKOMAN CORP.

DR WPI; 1996-015269/02.

XX New mutant streptavidin (SA) gene - fused to a firefly luciferase

PT gene, for the recombinant prepn. of a SA-FL fused protein

PS Disclosure; Page 10; 12pp; Japanese.

XX The present sequence is the L. lateralis luciferase, which can be

CC used in the construction of a streptavidin/luciferase fusion

CC protein, comprising a mutant Streptomyces avidinii streptavidin

CC and the L. lateralis (firefly) luciferase. The fusion protein

CC can be prep. by inserting the recombinant DNA encoding it into a

CC Escherichia species microbe, culturing the transformed microbe in a

CC medium and collecting the fusion protein from the culture. The

CC fusion protein can be used in industry.

XX Sequence 548 AA:

Query Match 99.9%; Score 2819; DB 17; Length 548;

Best Local Similarity 99.8%; Pred. No. 2.5e-235;

Matches 547; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MEMMENDENIVGPEPEFYPIEGSAGAO LRKYMRYAKGATFTNALTGVDYTYAEYLE 60

Db 1 memmendenivgpepfyleegsagaqlrkymdryaklgatftnaltdydyaeyle 60

Oy 61 KSCCIGALKNYGLVVDGRALCSENCCEFFIPVLGLFVGAVPAPTNELTYLRELVHSL 120

Db 61 kscclgealnknyglvvdgralcsenceeffipvlaglfgvavaptneltylrelvhs 120

Oy 121 GISKPTIVSSKKGLDKVITVOKTVAIKTIVLDSKVDRGQSDNFTKKKTPGCGFKG 180

Db 121 giskptivsskkgldkvitvoktvtalktivildskvdrygsmdnftkktpgqfkx 180

Oy 181 SSFTVVRKKEQVALIMNSSGTLPGVQLTHENLVTFSSHARPPIYGNQVSPGTAIL 240

Db 181 ssftvevnrkeqvalimnssgscglpkyqlthenlvtfsharppiyngnvspgtail 240

Oy 241 TVVFNHFGFMFTTLGYLTGCFRIYVMTLKFDEETFLKTDYKCCSVILVPLFAILNRS 300
 Db 241 tvvpfnhfgfmfttlgyltcgfrivmltkfdeetflktldykccsvllvplfai 300
 Oy 301 ELIDKYDLSNLVEIASGAPLSKEIGEAARPNLPGVNQGIGLETTTSATIIITTEBGDK 360
 Db 301 elldkydlsnlveiasgaplskeigearrpnlpgvnqygiletttsailltpegd 360
 Oy 361 PGASGVVPLFKAKVIDLDLTKTLPNRRGECVKGPMIMKGVDPNPEATREIIEEGWL 420
 Db 361 pgasgvvplfkakvidldtkclpnrtrgevcvkgpmlmkgyvdpnpeatreiide 420
 Oy 421 HTGDIGYDEEKHFFIVDRLSLIRKYGQVPAPAELESVLLHPNIFDAGVAVDPDPIAG 480
 Db 421 htgdigydeekhffivdrkslikygyvppaelesvllghpnifdagvavdpdp 480
 Oy 481 ELPGAVVVLKKGKSMTEKEVMDYVAASQVSNAKRLRGVRFVDPVPRKGLKIDGAKAREI 540
 Db 481 elpgavvvllkkgksmtekevmdyvaasvsnakrlrgvrtivdevpkqltqkldgkaltrei 540
 Oy 541 LKKPVAKM 548
 Db 541 lkkpvakm 548

RESULT 4

AAW53884 ID AAW53884 standard; Protein: 548 AA.

AC AAW53884;

DT 16-JUL-1998 (first entry)

DE Antibody-Firefly Luciferase fusion protein.

KM Firefly luciferase; antibody-luciferase fusion protein.

OS Synthetic.

PN JP09187281-A.

PD 22-JUL-1997.

PF 09-JAN-1996; 96JP-0001812.

PR 09-JAN-1996; 96JP-0001812.

PA (KIKK) KIKKOMAN CORP.

DR WPI; 1998-275089/25.

XX Antibody-firefly luciferase fused protein - and related products

PT i.e. firefly luciferase fused gene, recombinant DNA and its

PS preparation

XX Disclosure; Page 9-10; 17pp; Japanese.

CC This sequence is a fusion protein of the invention. The protein is

CC a antibody-firefly luciferase fusion protein, in which an antibody part

CC consisting of a peptide having antibody activity is combined with an

CC enzyme part consisting of firefly luciferase.

XX Sequence 548 AA:

Query Match 99.9%; Score 2819; DB 19; Length 548;

Best Local Similarity 99.8%; Pred. No. 2.5e-235;

Matches 547; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MEMMENDENIVGPEPEFYPIEGSAGAO LRKYMRYAKGATFTNALTGVDYTYAEYLE 60

Db 1 memmendenivgpepfyleegsagaqlrkymdryaklgatftnaltdydyaeyle 60

```

QY 61 KSCCIGELKNTGLVVDGRIALCSENCEEFFIPVLAGFIGVAPNTEIYTLRELIVSL 120
DB 61 kscclgealknryglvvdgrialscenceeffipvlagfigvaptneiytlrelivsl 120
QY 121 GISKPTIVSSKKGIDKYITVQKTYTAKITVIILDSKVDYRGYSMDNFIKKNTPOGFKG 180
DB 121 giskptivsskkgidkyitvqktytakitviiildskvdyrgysmdnfikknppgfk 180
QY 181 SSFKTEVNRKEQVALIMNSSGSTGLPKGYOLTHENLVTRFSHARDPIYGNQVSPGTA 240
DB 181 ssfktevnrkeqvalimnssgstglpkgyolthenlvtrshardpiygnvsgpta 240
QY 241 TVPPEHGEFMTTGLYITCGFRIVMLTKPDEEFLKTLQDYKCSSVILVPTLFAILNRS 300
DB 241 tvpphgefmgttglyitcgfrivmltkpdeeflktlodykcssvilvptlfaillnrs 300
QY 301 ELLDKYDLSNIVETASGAPLSKEIGEAVARFNLPGVROGYGLTETSATITTPBEDDK 360
DB 301 elldkydlisnlvelasgapslskeigeavarfnlpgvrgyglteetsailltpbeddk 360
QY 361 PGASGKYVPLFKAKAVIDLTKTTLGPNRGEVCVKGPMKMGVYDNPDEATREIIDEEGWL 420
DB 361 pgasgkyvplfkakavidltdtktlgpnrgvecvkgpmlmkgyvdpdeatreiideegwl 420
QY 421 HTGDIgyDEEKHFFIYDRLKSLIKYGYOPPALESVILQHNITPDAGVAPDPPIAG 480
DB 421 htgdigydeekhfifidrlkslikygyvppaelasvillqhnitpdagvavpdpbiag 480
QY 481 ELPGA VVVLKKGKSMTEKEVMDVYASQVSNAKRLRGVRFVDEVPKGLTGKIDGKAIR 540
DB 481 elpgavvvllkkgksmtekevmdyvasqvsnakrlrgsvrfvdevpkyglgkdgkair 540
QY 541 LKRPVAKM 548
DB 541 lkkrpvakm 548

```

RESULT 5
ID AAM99367 standard; Protein; 548 AA.
XX AAM99367;
AC AAM99367;
XX
DT 21-MAY-1999 (first entry)
XX
DE Mutant firefly luciferase protein #4.
XX
KM Bioluminescent protein; catalytic efficiency; stability; firefly;
KW luciferase; chimeric; luciola cruciata; luciola lateralis;
KW Photinus pyralis; primer; PCR; amplification.
XX
OS Luciola lateralis.
OS Synthetic.
XX
PN WO9902697-A1.
XX
PD 21-JAN-1999.
XX
PF 30-JUN-1998; 98WO-JP02936.
XX
PR 08-JUL-1997; 97US-0051917.
XX
PA (KIKK) KIKKOMAN CORP.
XX
PI Hirokawa K, Kajiyama N, Murekami S;
XX
DR WPI; 1999-120898/10.
XX
DR N-PSDB; AAX25718.
XX
PT New bioluminescent protein with improved properties - has greater
PT catalytic efficiency and stability and is obtained by modification
PT of natural precursors

```

XX PS Example 5; Page 36-39; 53pp; Japanese.
XX CC
CC The invention relates to the generation of bioluminescent proteins with
CC improved catalytic efficiency and stability. The proteins are generated
CC by addition, deletion or substitution of amino acids from a precursor
CC protein e.g. firefly luciferase, or by constructing chimeric luciferase
CC proteins from luciferase genes from e.g. Luciola cruciata, L. lateralis
CC and Photinus pyralis. This sequence represents an example of a mutated
CC luciferase protein of the invention.
XX CC
XX Sequence 548 AA:
SQ

```

Query Match. 99.8%; Score 2817; DB 20; Length 548;
Best Local Similarity 99.8%; Pred. No. 3,8e-235;
Matches 547; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MEMMENDENIVYGPPEFYPIERGSAQOLRYKMDRYAKLGAIFATNALTCVDYTYAEYLE 60
DB 1 memmendeniivyppefypieergsaqlrykmdryaklgaifatalnltgydytyaeyle 60
QY 61 KSCCIGELKNTGLVVDGRIALCSENCEEFFIPVLAGFIGVAPNTEIYTLRELIVSL 120
DB 61 kscclgealknryglvvdgrialscenceeffipvlagfigvaptneiytlrelivsl 120
QY 121 GISKPTIVSSKKGIDKYITVQKTYTAKITVIILDSKVDYRGYSMDNFIKKNTPOGFKG 180
DB 121 giskptivsskkgidkyitvqktytakitviiildskvdyrgysmdnfikknppgfk 180
QY 181 SSFKTEVNRKEQVALIMNSSGSTGLPKGYOLTHENLVTRFSHARDPIYGNQVSPGTA 240
DB 181 ssfktevnrkeqvalimnssgstglpkgyolthenlvtrshardpiygnvsgpta 240
QY 241 TVPPEHGEFMTTGLYITCGFRIVMLTKPDEEFLKTLQDYKCSSVILVPTLFAILNRS 300
DB 241 tvpphgefmgttglyitcgfrivmltkpdeeflktlodykcssvilvptlfaillnrs 300
QY 301 ELLDKYDLSNIVETASGAPLSKEIGEAVARFNLPGVROGYGLTETSATITTPBEDDK 360
DB 301 elldkydlisnlvelasgapslskeigeavarfnlpgvrgyglteetsailltpbeddk 360
QY 361 PGASGKYVPLFKAKAVIDLTKTTLGPNRGEVCVKGPMKMGVYDNPDEATREIIDEEGWL 420
DB 361 pgasgkyvplfkakavidltdtktlgpnrgvecvkgpmlmkgyvdpdeatreiideegwl 420
QY 421 HTGDIgyDEEKHFFIYDRLKSLIKYGYOPPALESVILQHNITPDAGVAPDPPIAG 480
DB 421 htgdigydeekhfifidrlkslikygyvppaelasvillqhnitpdagvavpdpbiag 480
QY 481 ELPGA VVVLKKGKSMTEKEVMDVYASQVSNAKRLRGVRFVDEVPKGLTGKIDGKAIR 540
DB 481 elpgavvvllkkgksmtekevmdyvasqvsnakrlrgsvrfvdevpkyglgkdgkair 540
QY 541 LKRPVAKM 548
DB 541 lkkrpvakm 548

```

RESULT 6
ID AAR30803 standard; Protein; 548 AA.
XX AAR30803;
XX
AC AAR30803;
XX
DT 24-MAY-1993 (first entry)
XX
DE Thermostable luciferase (wild-type).
XX
DE Firefly; Luciola cruciata; GENJ1; Luciola lateralis; HEIKE;
XX
KW mutant; assay.
XX
OS Luciola lateralis.

```

XX Key Location/Qualifiers
FH Misc-difference 217
FT /note- "amino acid replaced by hydrophobic amino
FT acid, esp. Ile, Leu or Val, for heat stable
FT mutant"
XX
XX EP524448-A.
XX
XX 27-JAN-1993.
XX
XX 26-JUN-1992; 92EP-0110808.
XX
XX 27-JUN-1991; 91JP-0157117.
XX 29-NOV-1991; 91JP-0317064.
XX 22-MAY-1992; 92JP-0131057.
XX
XX (KIRK ) KIRKMAN CORP.
XX
XX Elchi N, Naoki K;
XX
XX WPI; 1993-028553/04.
XX N-PSDB; AAQ34745.
XX
XX New DNA sequence of a thermostable luciferase - has the aminoacid
XX at position-217 of Luciola cruciata or lateralis luciferase
XX replaced by a hydrophobic aminoacid
XX
XX Claim 10-12; Page 30 + 25-29; 33pp; English.
XX
XX A thermostable luciferase of a firefly, in which an amino acid
XX at the 217 position or an amino acid at the position equiv. to the
XX 217 position of the luciferase of Luciola cruciata (GENJI firefly)
XX of Luciola lateralis (HEIKE firefly) is replaced by a hydrophobic
XX amino acids, esp. Ile, Leu or Val, is claimed. The mutant luciferase
XX is identical in properties to the wild-type luciferase except that
XX it is stable when heated to high temps., e.g. 50 degrees C.
XX The luciferase can be used in assays for e.g. ATP.
XX
XX Sequence 548 AA:
SQ
Query Match 99.7%; Score 2814; DB 14; Length 548;
Best Local Similarity 99.6%; Pred. No. 6.9e-235;
Matches 546; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 MENMENDENIYGPPEFYPIEGSAGAOIRKYMRYAKIGAIATFNALTGVDYTAETYLE 60
XX
XX 1 memmendeniygpepfyleegsagaqlrkymdryakigalaifnalgtvgdytaeyle 60
XX
XX 61 KSCCLGALKNKGIVVDGRIALCSENCPEFIPVLAGLFGVAPRTNEITLRELVSIL 120
XX
XX 61 kscclgealknkgivvdgrialsenceefipvlaglfigvaprtnetlrelvshl 120
XX
XX 121 GIKSPITVSSKGLDKVTIVOKTVAIKTIVILDSKVDYRGYSMDNEIKNTPGFRG 180
XX
XX 121 giskpitvsskglkvtivoktvaiktiivildskvdyrgysmdneiknptpgfkg 180
XX
XX 121 giskpitvsskglkvtivoktvaiktiivildskvdyrgysmdneiknptpgfkg 180
XX
XX 181 SSRKVEVNRKEQVALIMNSGSGTGPKYQVLTHENLVYRFSNARPIYGNQVSPETALL 240
XX
XX 181 ssrkvevnrkeqvalimnsstgtgpkylthenvlvrfsnarpiygnqvspetall 240
XX
XX 181 ssrkvevnrkeqvalimnsstgtgpkylthenvlvrfsnarpiygnqvspetall 240
XX
XX 241 TVVPFHNGFGEMFTLVTCGFRIVMLTKRDEEFLTKLQDYCCSVIIVPTLFAILNLS 300
XX
XX 241 tvvpfhngfgemftlvtcgfrivmltkrdeefltkldyccsviivptlfaillns 300
XX
XX 241 tvvpfhngfgemftlvtcgfrivmltkrdeefltkldyccsviivptlfaillns 300
XX
XX 301 ELIDKYLNLVIEIASGAPLSEIGEAARPNLPQVRGYSGLTETTSALITTPSGDK 360
XX
XX 301 elldkylnlvieiasgaplseigearpnlpqvrqysgltettsalittpsgdk 360
XX
XX 301 elldkylnlvieiasgaplseigearpnlpqvrqysgltettsalittpsgdk 360
XX
XX 361 PGASGKVVPLFKAKVIDLDKTKTLGPRRGEVCKGPMILKKGVDNDEARREIIDEQNL 420
XX
XX 361 pgasgkvvplfkakvidldkktlgnrtgvecvkqpmilmkgyvndpearetreideqnl 420

```

```

XX
XX 421 HTGDICGYDEEKHFIVDRLSLIKXGYVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
XX
XX 421 htgdicgydeekhfivdrlslikxyvppaelesvllqhpnidagvavdpdiag 480
XX
XX 481 ELPGAVVVLKKGSKMTEKEVMYVASQVSNAKRLRGVRFVDEVPKGLTGKIDKAIREI 540
XX
XX 481 elpgavvvllkkgksmtekevmvyasqvsnakrlrgvrfvdevpkgltgkidairei 540
XX
XX 541 LKRPVAKM 548
XX
XX 541 lkpvakm 548
XX
XX
XX
XX RESULT 7
XX AAR33710
XX ID AAR33710 standard; Protein; 548 AA.
XX
XX AAR33710;
XX
XX 24-MAY-1993 (first entry)
XX
XX DE Thermostable luciferase (T217L,L,V).
XX
XX DE Firefly; Luciola cruciata; GENJI; Luciola lateralis; HEIKE;
XX KM mutant; assay.
XX
XX OS Luciola lateralis.
XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 217
XX FT /label= ILE, LEU, VAL
XX
XX EP524448-A.
XX
XX 27-JAN-1993.
XX
XX 26-JUN-1992; 92EP-0110808.
XX
XX 27-JUN-1991; 91JP-0157117.
XX 29-NOV-1991; 91JP-0317064.
XX 22-MAY-1992; 92JP-0131057.
XX
XX (KIRK ) KIRKMAN CORP.
XX
XX Elchi N, Naoki K;
XX
XX WPI; 1993-028553/04.
XX
XX New DNA sequence of a thermostable luciferase - has the aminoacid
XX at position-217 of Luciola cruciata or lateralis luciferase
XX replaced by a hydrophobic aminoacid
XX
XX Claim 10-12; Page 30 + 25-29; 33pp; English.
XX
XX A thermostable luciferase of a firefly, in which an amino acid
XX at the 217 position or an amino acid at the position equiv. to the
XX 217 position of the luciferase of Luciola cruciata (GENJI firefly)
XX of Luciola lateralis (HEIKE firefly) is replaced by a hydrophobic
XX amino acids, esp. Ile, Leu or Val, is claimed. The mutant luciferase
XX is identical in properties to the wild-type luciferase except that
XX it is stable when heated to high temps., e.g. 50 degrees C.
XX The luciferase can be used in assays for e.g. ATP.
XX
XX Sequence 548 AA:
SQ
Query Match 99.7%; Score 2814; DB 14; Length 548;
Best Local Similarity 99.6%; Pred. No. 6.9e-235;
Matches 546; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 MENMENDENIYGPPEFYPIEGSAGAOIRKYMRYAKIGAIATFNALTGVDYTAETYLE 60
XX
XX 1 memmendeniygpepfyleegsagaqlrkymdryakigalaifnalgtvgdytaeyle 60

```

Db 1 memendeniyypepfypleegsagaqlrkymdryaklgalaftnaltygvdyaeyle 60
 QY 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFFIPVLAGFTIGVAPPTNEITRLVHSL 120
 Db 61 kscclgealknyglvvdgrialscenceeffipvlagftigvaptneitrlvhs 120
 QY 121 GISKPTIVFSSKKGLDKVITYOKVTAIKTVILDSKVDRGYOSMDNFIRKNTPOGFKG 180
 Db 121 giskptivfsskglkdvityokvtaiktvildskvdrgyosmdnfirkntpogfk 180
 QY 181 SSFKTEVNRKEQVALIMNSGSTGLPKGVOLTHENLVTRSHARDPIYGNQVSPGTAIL 240
 Db 181 ssfktevnrkeqvalimnsgstglpkgyvlthenvtrshardpiygnvspgtail 240
 QY 241 TYVPEHHGFMFTTLYGLTCGFRIVMLTKFDETFELKTLQDYKSSVILVPTLFAILNRS 300
 Db 241 tyvpfhngfmitllylctcgfrivmltkfdeetfkltdqykcsvllvptlfaillrs 300
 QY 301 ELLDKYDLSNVEIASGAPLSKEIGEAVARFNLPGVRQGYGLTETTSAILITPEGDDK 360
 Db 301 elldkydlslnveiasgapslskeigeavarfnlpgvrqgygltettsailitpegddk 360
 QY 361 PGASGKVVPLFKAKVIDLDRKTLGPNRGECVKGPMLMKGVDNDEATREIIDEQWL 420
 Db 361 pgasgkvvplfkakvidldrktlgnprgevcvkgpmlmkgyvndeatreideegwl 420
 QY 421 HTGDIGYDEEKHFPTVDRKLSLTKYGVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
 Db 421 htgdigydeekhfivdrklslikkygvpvpaesvllqhpnidagvagvdpdpiag 480
 QY 481 ELPGAVVVLKKGSMTEKEVMDYVASOVSNARKRLRGVRVDEVPKGLTKTGKAIKREI 540
 Db 481 elpgavvvleksgmtetekemdyvasovsnarkrlrgvrvdevpkgltktgkaiirei 540
 QY 541 LKKPVAKM 548
 Db 541 lkkpvakm 548
 Db 541 lkkpvakm 548
 RESULT 8
 AAM12395
 ID AAM12395 standard; Protein; 636 AA.
 AC AAM12395;
 DT 20-MAY-1997 (first entry)
 DE Firefly luciferase conjugated with 87 aa biotin peptide at 3' end.
 XX Fusion protein: firefly; luciola lateralis; luciferase; biotinylation;
 KM wild type; E.coli; bioluminescence assay; quantification; ligand;
 KM receptor.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Region 1..547
 FT Region /note="Firefly luciferase portion of fusion protein"
 FT Region 548..549
 FT /note="encoded by linker sequence used to construct
 FT Region fusion gene"
 FT Region 549..636
 FT /note="biotin peptide portion of fusion protein"
 XX JP08308578-A.
 XX
 PD 26-NOV-1996.
 XX
 PD 24-APR-1995; 95JP-0098857.
 XX
 PR 14-MAR-1995; 95JP-0054625.
 PR 27-JUL-1994; 94JP-0193798.
 XX

PA (KIKK) KIKKOMAN CORP.
 XX
 DR WPI: 1997-059697/06.
 DR N-PSDB; AAT63269.
 XX
 PT Fusion protein comprising firefly luciferase and biotinylated
 PT peptide - useful in a bio-luminescent analytical method for
 PT quantifying ligands
 XX
 PS Claim 2; Page 12-13; 13pp; Japanese.
 XX
 CC This is the amino acid sequence of a novel fusion protein which comprises
 CC the firefly (Luciola lateralis) luciferase protein and an 87 amino acid
 CC biotinylated peptide designated the E.coli biotin carboxy-carrier protein
 CC (BCCP-87). The fusion gene was generated by firstly amplifying the
 CC sequence encoding the E.coli BCCP-87 from the E.coli genome and inserting
 CC the resultant sequence into the 3' end of the firefly luciferase gene in
 CC plasmid pHLf230. The biotin peptide sequence replaces the C-terminal Met
 CC residue of the wild type luciferase sequence. The plasmid was then
 CC introduced into E.coli JM101 for production of the fusion protein. The
 CC luciferase ligands which may modulate binding of luciferase to its
 CC receptor.
 CC
 CC Sequence 636 AA;
 XX
 SQ

Query Match 99.7%; Score 2814; DB 18; Length 636;
 Best Local Similarity 99.8%; Pred. No. 8, 6e-235;
 Matches 546; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEMNENDENIYVPEPFYPLEEGSAGAQLRKYMDRYAKLGALFNTNALNGVDYAEYLE 60
 Db 1 memendeniyypepfypleegsagaqlrkymdryaklgalaftnaltygvdyaeyle 60
 QY 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFFIPVLAGFTIGVAPPTNEITRLVHSL 120
 Db 61 kscclgealknyglvvdgrialscenceeffipvlagftigvaptneitrlvhs 120
 QY 121 GISKPTIVFSSKKGLDKVITYOKVTAIKTVILDSKVDRGYOSMDNFIRKNTPOGFKG 180
 Db 121 giskptivfsskglkdvityokvtaiktvildskvdrgyosmdnfirkntpogfk 180
 QY 181 SSFKTEVNRKEQVALIMNSGSTGLPKGVOLTHENLVTRSHARDPIYGNQVSPGTAIL 240
 Db 181 ssfktevnrkeqvalimnsgstglpkgyvlthenvtrshardpiygnvspgtail 240
 QY 241 TYVPEHHGFMFTTLYGLTCGFRIVMLTKFDETFELKTLQDYKSSVILVPTLFAILNRS 300
 Db 241 tyvpfhngfmitllylctcgfrivmltkfdeetfkltdqykcsvllvptlfaillrs 300
 QY 301 ELLDKYDLSNVEIASGAPLSKEIGEAVARFNLPGVRQGYGLTETTSAILITPEGDDK 360
 Db 301 elldkydlslnveiasgapslskeigeavarfnlpgvrqgygltettsailitpegddk 360
 QY 361 PGASGKVVPLFKAKVIDLDRKTLGPNRGECVKGPMLMKGVDNDEATREIIDEQWL 420
 Db 361 pgasgkvvplfkakvidldrktlgnprgevcvkgpmlmkgyvndeatreideegwl 420
 QY 421 HTGDIGYDEEKHFPTVDRKLSLTKYGVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
 Db 421 htgdigydeekhfivdrklslikkygvpvpaesvllqhpnidagvagvdpdpiag 480
 QY 481 ELPGAVVVLKKGSMTEKEVMDYVASOVSNARKRLRGVRVDEVPKGLTKTGKAIKREI 540
 Db 481 elpgavvvleksgmtetekemdyvasovsnarkrlrgvrvdevpkgltktgkaiirei 540
 QY 541 LKKPVAK 547
 Db 541 lkkpvak 547

RESULT 9

AAW48855
ID AAW48855 standard; protein; 640 AA.
AC AAW48855;
XX
XX 10-SEP-1998 (first entry)
DT
XX Luciferase-lyso-staphin fused protein.
DE Luciferase-lyso-staphin fused protein.
XX Luciferase-lyso-staphin fused protein; chimeric; bioluminescence.
KW
XX Chimeric - Staphylococcus simulans.
OS Chimeric - Luciola lateralis.
XX
XX Key Location/Qualifiers
FH Domain 1..549
FT /note= "Luciferase protein"
FT 550..640
FT /note= "Lyso-staphin protein"
FT
XX JP10150991-A.
XX
XX 09-JUN-1998.
XX
XX 25-NOV-1996; 96JP-0328042.
XX
XX 25-NOV-1996; 96JP-0328042.
XX
XX (KIKK) KIKKOMAN CORP.
XX
XX WPI: 1998-379994/33.
XX N-PSDB; AAV32467.
XX
XX New luciferase-lyso-staphin fused protein - useful for
PT bioluminescence analysis
PS
XX Claim 4; Pages 7-9; 10pp; Japanese.
XX
XX The invention claims for a luciferase-lyso-staphin fused gene which
CC encodes a fused protein in which a peptide part consisting of
CC luciferase, from Luciola lateralis, is connected to a peptide part
CC consisting of lyso-staphin from Staphylococcus simulans. The method
CC described by the invention can be used to prepare a luciferase-
CC lyso-staphin fused protein efficiently. The fused protein is useful
CC for bioluminescence analysis.
XX
XX Sequence 640 AA:
SQ

Query Match 99.7%; Score 2814; DB 19; Length 640;
Best Local Similarity 99.8%; Pred. No. 8.7e-235;
Matches 546; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MENMENDENIYVGPPEPIEESGAGOLRKYMDRYAKLGAIAFTNALTGVDYTAEYLE 60
DB 1 memmendeniivgpepfyieesagaglrkymdryaklgatlnatlgvdytaeyle 60
QY 61 KSCCCEALKNGLVVDGRALCSECEEFIPVLAGLFGVAPATNEIYTLRELVS 120
DB 61 kscclgealknglvvdgralcsenceefipvlaglfgvapatneiytlrelvshl 120
QY 121 GISKPTIVSSKKGDLKVTVOCTVTAIKTIYILDSKVDYRGYQSDMNFIKNTPOGFG 180
DB 121 giskptivsskkgdlkvtvocktvtaiktivildskvdrygysmdnfikntlpqgfk 180
QY 181 SSKFTEVNRKEQVALIMNMSGTGLPKGVOLTHENLVTRFSHARDPITGNVSPETAL 240
DB 181 sskftevnrkeqvalimnmsgtgllpkgvollhenlvtrfsharpiygnvsgpetal 240
QY 241 TVVPEHNGPMFTTIGTCGFRIVMLTKRDEFTPLKTLDDYKSSVIVPTLFFALINS 300
DB 241 tvvpfhngpmfttltgctgfriivmltkrdeetflkldgkssvllvptlffalins 300

QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARBPNLPGVROGYLTETTSALITPEGD 360
DB 301 elldkydlslnlveiasgagplskelgeavarfnlpvrgygltsaliltpgedk 360
QY 361 PGASGRVPLFKAKVIDLDTKKTGPNRGEVCVGPMLMKGVNDNPEATREIIDE 420
DB 361 pgasgrvplfkakvidldtkktlgnrrgevcvkgpmlmkgyvndnpeatreiide 420
QY 421 HTGDDIGYDEEKHFIVDLKSLIKYKGYVPPAELESVLLQHPNIFDAGVGPDP 480
DB 421 htgdigydeekhfivdlkslikkygvppaellesvllqhpnidagvgpdpd 480
QY 481 ELPGAVVLEKKGSKMTEKEVMDYVASOVSNAKRLRGVRFVDEVPKLGKIDGAI 540
DB 481 elpgavvlekgksmtekevmdyvasovsnakrlrgvrfvdevpklgkidgair 540
QY 541 LKRPVAK 547
DB 541 lkrpvak 547

RESULT 10
AAR03731
ID AAR03731 standard; protein; 548 AA.
XX
XX AAR03731;
XX
XX 22-JUL-1990 (first entry)
XX
XX
XX Luciola lateralis recombinant luciferase.
XX
XX Luciferase; enzyme; ATP; plasmid pHLf7; firefly.
XX
XX Luciola lateralis.
XX
XX EP353464-A.
XX
XX 07-FEB-1990.
XX
XX 30-JUN-1989; 89EP-0111958.
XX
XX 22-DEC-1988; 88JP-0162402.
XX 01-JUL-1988; 88JP-0162402.
XX
XX (KIKK) KIKKOMAN CORP.
XX
XX
XX Tatsumi H, Kajiyama N, Nakano E;
XX
XX WPI: 1990-038240/06.
XX N-PSDB; AAQ03257.
XX
XX New gene encoding luciferase -
PT derived from Luciola lateralis and expressed in and purified
PT from Escherichia coli genus cells.
XX
XX Claim 2; page 21; 43pp; English.
XX
XX This enzyme is useful for determining ATP levels. It is
CC contained in plasmid pHLf7.
CC
XX Sequence 548 AA:
SQ

Query Match 99.6%; Score 2811; DB 11; Length 548;
Best Local Similarity 99.5%; Pred. No. 1.2e-234;
Matches 545; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MENMENDENIYVGPPEPIEESGAGOLRKYMDRYAKLGAIAFTNALTGVDYTAEYLE 60
DB 1 memmendeniivgpepfyieesagaglrkymdryaklgatlnatlgvdytaeyle 60
QY 61 KSCCCEALKNGLVVDGRALCSECEEFIPVLAGLFGVAPATNEIYTLRELVS 120
DB 61 kscclgealknglvvdgralcsenceefipvlaglfgvapatneiytlrelvshl 120

```

Db 61 kscclgealknyjlvdgrialcsenceeffipvlaglflygvavptneylirelvhs1 120
QY 121 GISRPTEVSSKRGIDKVTVOKTVAIKTIVILDSKVYGYOSMNEFKKNTPOGFKG 180
Db 121 gisrptvsskrgidkvtvoktvaiktivildskvdygyosmndfikkntpgfkg 180
QY 181 SSFRTVEVNRKEQVALIMNSSGSGLPKGVOLTHEMLVTRFSHARDPIYGNVSPGTAIL 240
Db 181 ssftrvevnrkeqvalimnssgstglpkgyvltthenavtrfshardpiygnvsgptail 240
QY 241 TVPVEHHGFGMFTTIGLTCGFRIVMLTKPDEEFLTKTDYKCSSYILVPTLFAILNRS 300
Db 241 tvpvehhgfgmfttliglctcgrivmltkpdeefltktldykcssyilvptlfailnrs 300
QY 301 ELDDKYLNLVETASGAPLSKEIGAVARRFNLPGVRGOGYGLTETTSAILITPEGDGK 360
Db 301 elldkylnlvetaasgapslskeigavarrrfnlpgvrsgygltettsailitpegdck 360
QY 361 PGASGVVPLFRKAVIDLTDTKKTIGPVRGECVCKGPMLMKGYVNDPEATREITIDEGWL 420
Db 361 pgasgvvplfrkavidldtktktlgnprirgevcvkpmlmkgyvndpeatreitideegwl 420
QY 421 HTGDIGYDEEKHFIVDRLSLIKRYGVOPPALESVLLQHPNIFDAGVAGVPDPIAG 480
Db 421 htgdigydeekhfivdrllsklikrygvoppaesvllqhpnidagvagvpdpiaag 480
QY 481 ELRGAVVVLKKGSMTEKEMDYVASQVSNARKLRGVRFVDEVPKGLTGKIDGKAIREI 540
Db 481 elrgavvvllkkgsmtekevmdyvasqvsnakrlrgvrfvdelpkgltgkldgkaiirei 540
QY 541 LKKPVAKM 548
Db 541 lkkpvakm 548

RESULT 11
AAW12394
ID AAW12394 standard; Protein; 568 AA.
AC
XX AAW12394;
XX
DT 20-MAY-1997 (first entry)
DE Firefly luciferase conjugated with 23 aa biotin peptide at N-terminus.
XX
KM Fusion protein: firefly: luciola lateralis; luciferase; biotinylation;
KW wild type; E.coli; bioluminescence assay; quantification; ligand;
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..24
FT Region /note= "biotin peptide"
FT Region 25..568
FT Region /note= "firefly luciferase"
XX
PN JP08308578-A.
PD 26-NOV-1996.
XX
PE 24-APR-1995; 95JP-0098857.
PR 14-MAR-1995; 95JP-0054625.
PR 27-JUL-1994; 94JP-0193798.
XX
PA (KIKK ) KIKKOMAN CORP.
DR WPI: 1997-059697/06.
DR N-PSDB: AAT63268.
XX
PT Fusion protein comprising firefly luciferase and biotinylated
PT peptide - useful in a bio-luminescent analytical method for

```

```

PT quantifying ligands
XX
PS Claim 2: Page 10-11; 13pp; Japanese.
XX
CC This is the amino acid sequence of a novel fusion protein comprising
CC the firefly (luciola lateralis) luciferase protein and a 23 amino acid
CC biotinylated peptide designated biotin peptide #84. The gene was
CC generated by firstly annealing the sequence encoding peptide #84
CC (AAT63270) and its complementary sequence, then inserting the resultant
CC double stranded sequence into the 5' end of the luciferase gene in
CC plasmid pHL108. The biotin peptide sequence replaces the first 4 amino
CC acids of the wild type luciferase sequence. The plasmid was then
CC introduced into E.coli JM101 for production of the fusion protein. The
CC novel protein can be used in bioluminescence assays to quantify
CC luciferase ligands which may modulate binding of luciferase to its
CC receptor.
XX
SQ Sequence 568 AA:

```

```

Query Match 99.5%; Score 2801; DB 18; Length 568;
Best Local Similarity 99.5%; Pred. No. 9.7e-234;
Matches 543; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 NMENDENTVYGEPEPPYPIEESGAGALRKRYMDRYAKLGAIFTNALNGVDVTVAEYLEKS 62
Db 23 slendenlvypgepprypieesgsagqlrkymdryaklgaiftnalngvdytyaeyleks 82
QY 63 CCLGALNKGVLVNDGRALALSCENCEFFIPVLAGLFGVADPTNETITLRELVHSLGI 122
Db 83 cclgealnkgvlvndgrialalcsenceeffipvlaglflygvadptneylirelvhs1gl 142
QY 123 SKPTVFESSKRGIDKVTVOKTVAIKTIVILDSKVYRGYOSMNDTICKNTPOGFGSS 182
Db 143 skptlvfsskrgidkvtvoktvaiktivildskvdyrgyosmndfikkntpgfkgss 202
QY 183 FRTVEVNRKEQVALIMNSSGSGLPKGVOLTHEMLVTRFSHARDPIYGNVSPGTAILTV 242
Db 203 ftrvevnrkeqvalimnssgstglpkgyvltthenavtrfshardpiygnvsgptailtv 262
QY 243 VPFHHGFGMFTTIGLTCGFRIVMLTKPDEEFLTKTDYKCSSYILVPTLFAILNRS 302
Db 263 vpfhhgfgmfttliglctcgrivmltkpdeefltktldykcssyilvptlfailnrsel 322
QY 303 LDKYDLSNLVETASGAPLSKEIGAVARRFNLPGVRGOGYGLTETTSAILITPEGDGK 362
Db 323 ldkydlslnlvetaasgapslskeigavarrrfnlpgvrsgygltettsailitpegdckpg 382
QY 363 ASGVVPLFRKAVIDLTDTKKTIGPVRGECVCKGPMLMKGYVNDPEATREITIDEGWLHT 422
Db 383 asgvvplfrkavidldtktktlgnprirgevcvkpmlmkgyvndpeatreitideegwlht 442
QY 423 GDIGYDEEKHFIVDRLSLIKRYGVOPPALESVLLQHPNIFDAGVAGVPDPIAGEL 482
Db 443 gdigydeekhfivdrllsklikrygvoppaesvllqhpnidagvagvpdpiaagel 502
QY 483 PGAVVVLKKGSMTEKEMDYVASQVSNARKLRGVRFVDEVPKGLTGKIDGKAIREI 542
Db 503 pgavvvllkkgsmtekevmdyvasqvsnakrlrgvrfvdelpkgltgkldgkaiirei 562
QY 543 KPVAKM 548
Db 563 kpvakm 568

RESULT 12
AAW53883
ID AAW53883 standard; Protein; 568 AA.
AC
XX AAW53883;
XX
DT 16-JUL-1998 (first entry)
XX

```

DE Antibody-Firefly Luciferase fusion protein.
 XX Firefly luciferase; antibody-luciferase fusion protein.
 XX Synthetic.
 OS Luciola cruciata.
 XX JP09187281-A.
 PN 22-JUL-1997.
 XX 09-JAN-1996; 96JP-0001812.
 XX 09-JAN-1996; 96JP-0001812.
 PR 09-JAN-1996; 96JP-0001812.
 XX (KIKK) KIKKOMAN CORP.
 PA WPI; 1998-275089/25.
 XX N-PSDB; AAV23595.
 DR Antibody-firefly luciferase fused protein - and related products
 PT i.e. firefly luciferase fused gene, recombinant DNA and its
 PT preparation
 XX Disclosure; Page 11-13; 17pp; Japanese.
 PS This sequence is a fusion protein of the invention. The protein is
 CC an antibody-firefly luciferase fusion protein, in which an antibody part
 CC consisting of a peptide having antibody activity is combined with an
 CC enzyme part consisting of firefly luciferase.
 XX
 XX Sequence 568 AA:
 QY Query Match 99.2%; Score 2801; DB 19; Length 568;
 Best Local Similarity 99.5%; Pred. No. 9.7e-234;
 Matches 543; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 Db 3 NMENDEIVYGPEPPYPIEGSAGAOIRKYMRYAKLGAIAFTNALTGVDYTYAEYLEKS 62
 QY 23 slendeniivgpepfybiegsagqlrkymdryaklgatalfnaltgydytyaeyleks 82
 QY 63 CCGEALKNGVVDGRIALCSENCEEFIPVLAGLFTIGVGAFTMEITLRELTVHSLGI 122
 Db 83 cclgealkngylvvdgrialscenseceefipvlaglftigvgaftmeitlreltvhslgi 142
 QY 123 SKPTIVSSKKGLDKVITVOKTVAITKTIIVILDSKVDYRGYSMDNFIRKNTPOGFGSS 182
 Db 143 skptivsskkglkdvitvoktvtaktiivildskvdrygysmdnfirkntpqgfgkss 202
 QY 183 EKTVEVNRKEQVALIMSSSGSTGLPKVOULTHEMLVTRFSHARDPIYGNQVSPGTATLTV 242
 Db 203 fktvevnrkeqvalimsssgstglpkvoulthemlvtrfshardpiygnqvspgtatlty 262
 QY 243 VPFHNGFMFTTLYTCGFRIVMLTFKDEETFLKTLQDYKCSSVILVPTLFAILNSEL 302
 Db 263 vpfhngfmfttlytgcgfrivmltfkdeetflktdykcssvilvptlfaalnrsel 322
 QY 303 LDKYDLSNLVEIASGAPLSEKEIGENAVARRNLPGVQGVLTETTSALITTPGDDKPG 362
 Db 362 ldkydlslnlveiasgaplskeigeavartrnlpgvrgyglteetsaliltpegddkpg 382
 QY 383 ASGVVPLFKAFAKVIDLDRKKTGLPARRGECVKGPMLKKGVDNPEATREITDEGWLHT 422
 Db 422 asgvvplfkaakvidldrkkktglparrgecvkgpmlkkgvndppeatreideegwlht 442
 QY 442 GDIGYDEEKHEFIVDRILSKLIRKGYOVPAELESVILQHPNITFDACVAPPIAGEL 482
 Db 482 gdigydeekhefivdrilsklirkgyovpaelesvilqhpnitfdacvappiagel 502
 QY 502 PGAVVVLKKGKSGMTEKEVNDVYASOVNAKRLRGVRFDEVPKGLTKIDGAKIRELIK 542
 Db 542 pgavvvllkkgksgmtekevndvyasovnakrlrgvrfdevpkgltkidgakirelik 562

QY 543 KPVAKM 548
 Db 563 kvakm 568
 RESULT 13
 AAW04208
 ID AAW04208 standard; Protein; 673 AA.
 XX
 XX AAW04208;
 AC 12-NOV-1996 (first entry)
 XX
 XX Streptavidin/luciferase fusion protein.
 DE Streptavidin; luciferase; fusion protein;
 XX Streptomyces avidinii; Luciola lateralis; firefly;
 KW recombinant production; industry.
 KW
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH 1..128
 FT Protein /label= streptavidin
 FT 130..673
 FT Protein /label= luciferase
 XX
 XX JP07289264-A.
 PN 07-NOV-1995.
 XX 27-APR-1994; 94JP-0090275.
 XX 27-APR-1994; 94JP-0090275.
 PR 27-APR-1994; 94JP-0090275.
 XX (KIKK) KIKKOMAN CORP.
 PA WPI; 1996-015269/02.
 DR N-PSDB; AAT33850.
 XX New mutant streptavidin (SA) gene - fused to a firefly luciferase
 PT gene, for the recombinant prepn. of a SA-FL fused protein
 PT Example 1; Pages 6-7; 12pp; Japanese.
 XX
 XX The present sequence is a streptavidin/luciferase fusion
 CC protein, comprising the Streptomyces avidinii streptavidin gene
 CC and the Luciola lateralis (firefly) luciferase gene. The fusion
 CC protein can be prep'd by inserting the recombinant DNA encoding
 CC it into a Escherichia species microbe, culturing the transformed
 CC microbe in a medium and collecting the fusion protein from the
 CC culture. The fusion protein can be used in industry.
 XX
 XX Sequence 673 AA:
 QY Query Match 99.2%; Score 2801; DB 17; Length 673;
 Best Local Similarity 99.5%; Pred. No. 1.2e-233;
 Matches 543; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 Db 3 NMENDEIVYGPEPPYPIEGSAGAOIRKYMRYAKLGAIAFTNALTGVDYTYAEYLEKS 62
 QY 128 slendeniivgpepfybiegsagqlrkymdryaklgatalfnaltgydytyaeyleks 187
 QY 63 CCGEALKNGVVDGRIALCSENCEEFIPVLAGLFTIGVGAFTMEITLRELTVHSLGI 122
 Db 188 cclgealkngylvvdgrialscenseceefipvlaglftigvgaftmeitlreltvhslgi 247
 QY 123 SKPTIVSSKKGLDKVITVOKTVAITKTIIVILDSKVDYRGYSMDNFIRKNTPOGFGSS 182
 Db 248 skptivsskkglkdvitvoktvtaktiivildskvdrygysmdnfirkntpqgfgkss 307

QY 183 FKTVEVNRKEQVALIMNSSGSTGLPKGVOLTHENLVTRFSHARDPITYGNVSPGTAITLV 242
 Db 308 fktvevnrkeqvalimnssgstglpkgyqlthenlvtrfshardpitygnvspgtailtv 367
 QY 243 VPFHGFQMFETTLGTLGTFRIYMLTKFDEEFLKTLQDYKCSVILVPLFAILNSEL 302
 Db 368 vpfhgfqmfetltlgytlcgfrlvmktkideetflktldqkcsvillvplfailnssel 427
 QY 303 LDKYDLSNLVETIASGAPLSEKEIGEAVARPNLPGVRQGYLTETTSAILITPEGDDKPG 362
 Db 428 ldkydlslnlveliasgapslekeigearvfnlpgvrqgyltettsailitpegddkpg 487
 QY 363 ASGVVPLFKKAKVIDLDTKTLGNRRGECVCKGPMIMKGYVNDPEATRETIIDEGWLT 422
 Db 488 asgvvplfkakavidltdtktlgnrrgevcvkgpmlmkgvndpeatreliidegwlt 547
 QY 423 GDIGYDEEKHFIVDLKSLIKKGYVPPAELESVLLQHPNIFDAGVAGVDPPIAGEL 482
 Db 548 gdigydeekhfivdlkslikkyvppaellesvllqhpnidagvagvdpdpiagel 607
 QY 483 PGAVVVLKKGKSMTEKEVMDYVASQVSNARKLRGVRFPDEVKGLTGKIDGKAIREILK 542
 Db 608 pgavvvlekkgksmtekevmdyvasqvsnakrlrgvrfvdevpkgtlgtkldgkairreilk 667
 QY 543 KPYAKM 548
 Db 668 kpyakm 673

RESULT 14

AAW04209
 ID AAW04209 standard; Protein; 685 AA.

AC AAW04209;

DT 12-NOV-1996 (first entry)

DE Mutant streptavidin/luciferase fusion protein.

KW Mutant; streptavidin; luciferase; fusion protein;
 Streptomyces avidinii; Luciola lateralis; firefly;
 recombinant production; industry.

OS Synthetic.

FH Key Location/Qualifiers

FT Protein 1..140
 /label= mutant streptavidin

FT Protein 142..685
 /label= luciferase

FT Peptide 1..17
 /label= mutant N-terminal peptide

PN JP07289264-A.

XX 07-NOV-1995.

XX 27-APR-1994; 94JP-0090275.

XX 27-APR-1994; 94JP-0090275.

XX (RIKK) KIRKOMAN CORP.

XX WPI; 1996-015269/02.

XX N-PSDB; AAT33851.

XX New mutant streptavidin (SA) gene - fused to a firefly luciferase
 gene, for the recombinant prepn. of a SA-FL fused protein

XX Claim 4; Pages 8-9; 12pp; Japanese.

XX The present sequence is a mutant streptavidin/luciferase fusion
 protein, comprising a mutant Streptomyces avidinii streptavidin and

CC the Luciola lateralis (firefly) luciferase. The fusion protein can
 CC be prepd. by inserting the recombinant DNA encoding it into a
 CC Escherichia species microbe, culturing the transformed microbe in a
 CC medium and collecting the fusion protein from the culture. The
 CC fusion protein can be used in industry.
 SQ Sequence 685 AA;

Query Match 99.2%; Score 2801; DB 17; Length 685;
 Best Local Similarity 99.5%; Pred. No. 1.3e-233;
 Matches 543; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 NNDENNIYGEPEPYPIEESGAGOLRKYMRYAKLAIFTNALTGVDVTAIEYLS 62
 Db 140 slendeniyygepefpylieesagagllkymdryaklaiaftnalgvdytaieyels 199
 QY 63 CCLGEALKNYGLVNDGRALALSCNCEEFPIYLAGLFIGVGAAPNEIYTLRELVHSLGI 122
 Db 200 cclgealknvglvndgralalscnceeflpylaglfigvgaapneiytlrelvhsigl 259
 QY 123 SKPTIVFSSKKGLDKVITVOKTVAIKTIVILDSKVDRGYQSDMNFTKMTPOGFKGSS 182
 Db 260 skptivfsskkgldkvitvoktvaiktivildskvdyrgysmdnftkmtppgfkgyss 319
 QY 183 FKTVEVNRKEQVALIMNSSGSTGLPKGVOLTHENLVTRFSHARDPITYGNVSPGTAITLV 242
 Db 320 fktvevnrkeqvalimnssgstglpkgyqlthenlvtrfshardpitygnvspgtailtv 379
 QY 243 VPFHGFQMFETTLGTLGTFRIYMLTKFDEEFLKTLQDYKCSVILVPLFAILNSEL 302
 Db 380 vpfhgfqmfetltlgytlcgfrlvmktkideetflktldqkcsvillvplfailnssel 439
 QY 303 LDKYDLSNLVETIASGAPLSEKEIGEAVARPNLPGVRQGYLTETTSAILITPEGDDKPG 362
 Db 440 ldkydlslnlveliasgapslekeigearvfnlpgvrqgyltettsailitpegddkpg 499
 QY 363 ASGVVPLFKKAKVIDLDTKTLGNRRGECVCKGPMIMKGYVNDPEATRETIIDEGWLT 422
 Db 500 asgvvplfkakavidltdtktlgnrrgevcvkgpmlmkgvndpeatreliidegwlt 559
 QY 423 GDIGYDEEKHFIVDLKSLIKKGYVPPAELESVLLQHPNIFDAGVAGVDPPIAGEL 482
 Db 560 gdigydeekhfivdlkslikkyvppaellesvllqhpnidagvagvdpdpiagel 619
 QY 483 PGAVVVLKKGKSMTEKEVMDYVASQVSNARKLRGVRFPDEVKGLTGKIDGKAIREILK 542
 Db 620 pgavvvlekkgksmtekevmdyvasqvsnakrlrgvrfvdevpkgtlgtkldgkairreilk 679
 QY 543 KPYAKM 548
 Db 680 kpyakm 685

RESULT 15

AAW53882
 ID AAW53882 standard; Protein; 788 AA.

AC AAW53882;

DT 16-JUL-1998 (first entry)

DE Antibody-Firefly luciferase fusion protein.

XX Firefly luciferase; antibody-luciferase fusion protein.

XX Synthetic.

XX Luciola cruciata.

XX Key Location/Qualifiers
 FT Misc-difference 224 /note= "encoded by CAG"
 FT Misc-difference 247

FT /note= "encoded by GAT"
FT Misc-difference 457
FT /note= "encoded by TTG"

XX JP09187281-A.

XX 22-JUL-1997.

PF 09-JAN-1996; 96JP-0001812.

XX 09-JAN-1996; 96JP-0001812.

XX (KIKK) KIKKOMAN CORP.

XX WPI: 1998-275089/25.

DR N-PSDB; AAV23580.

XX Antibody-firefly luciferase fused protein - and related products
PT i.e. firefly luciferase fused gene, recombinant DNA and its
PT preparation

XX Claim 5; Page 14-16; 17pp; Japanese.

CC This sequence is the fusion protein of the invention. The protein is
CC a antibody-firefly luciferase fusion protein, in which an antibody part
CC consisting of a peptide having antibody activity is combined with an
CC enzyme part consisting of firefly luciferase.

XX Sequence 788 AA:

Query Match 98.7%; Score 2785; DB 19; Length 788;

Best Local Similarity 99.3%; Pred. No. 3.8e-232;
Matches 541; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 MENDENIYGPPEPYPYIEGSGAGAOIRKYMDRYAKIGAIPTNALGVDTYAEYLEKSC 63

Db :|||||
244 lenweniyygpepfypleegsagaqlrkymdryaklgaiafnaltgvdyaeyleksc 303

QY 64 CIGEAALKNYGLVNDGRIALCSNCEFFIPVLAGEFVGVAPTNEIYTRRELVSIGIS 123

Db :|||||
304 clgealknyglvndgrialcsenceefflvpvlaglligvavapneilylrelvmslgs 363

QY 124 KPTIVFSKKGLDKVTVVQKTVAITIKTIYILDSKVDYRGQSMDFIKKNTPOGFGSSSF 183

Db :|||||
364 kptivfskkglkdvltvqktvaktviltkskvdryrgysmdnfikntpqgfgsssf 423

QY 184 KTVENRRKEQVALIMNSGSGTGLPKGVQLTHENLVTRFSHARDPIYGNQVSPCTAALTIV 243

Db :|||||
424 ktvevnrkqevalimnsgstglprkvqlthenavtrfshardpiygnqvspqtaalltv 483

QY 244 PFHHGGMFTTGLYLCGFRIVMLTQDEBETFLKTQDYKSSVILPFLFALINRSELL 303

Db :|||||
484 pfhgmfttglylclgfrivmltkdeetflktqdykssvllpflfalinrsell 543

QY 304 DKYDLNLIWEIASGAPLSEKEIGEAVARRENLPQVROGYGLTETTSAILITPEGDDKPGA 363

Db :|||||
544 dkydlnliveiasgaplskeigearrenlpqvrqgyglteetsailltpgddkpga 603

QY 364 SGRVVPDLFAKVIDLDTKKTFLGPNRRGEVCKGPMIMKGYVDNPEATREI IDEEGWLHTG 423

Db :|||||
604 sgrvvpdlfkavidldtkktflgpnrrgevcvqpmimkgyvdnpeatreideegwlhtg 663

QY 424 DIGYDEEKHEFTVDRKLKSLIKKGYPAPLESVLLQHPNIFDAGVAGVPPDTAGELP 483

Db :|||||
664 digydeekheftvdrklkslikkgypaplesvllqhpnlfdagvavppdtagelp 723

QY 484 GAVVVLAKKSKMTEKEMVYVASQVSNARKLRGVRFDVEVPKGLTGKIDGKAIREILKK 543

Db :|||||
724 gavvvleksgmtekevmyvasqvsnakrlrgvrfvdevpkgltgkldgkairailkk 783

QY 544 PVAKM 548

|||||

Db 784 pvakm 788

Search completed: November 29, 2001, 01:53:12
Job time: 3252 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2001, 01:53:12 ; Search time 46.79 Seconds
(without alignments)
867.539 Million cell updates/sec

Title: US-09-581-241-6

Perfect score: 2623
Sequence: 1 MENNENDENTIVGPEPPYPI.....TGKIDGKAIREILKKPYAKM 548

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq.1101.*
1: /SID2/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2823	100.0	548	AA1980	L. lateralis lucif
2	2821	99.9	548	AA1981	L. lateralis lucif
3	2817	99.8	548	AA1982	Antibody-Firefly 1
4	2817	99.8	548	AA1983	Mutant firefly luc
5	2815	99.7	548	AA1984	Thermotable lucif
6	2814	99.7	548	AA1985	Thermotable lucif
7	2814	99.7	548	AA1986	Firefly luciferase
8	2812	99.6	548	AA1987	Luciferase-lysost
9	2812	99.6	548	AA1988	Luciola lateralis
10	2811	99.6	548	AA1989	Firefly luciferase
11	2799	99.1	548	AA1990	

12	2799	99.1	548	AA1991	Antibody-Firefly 1
13	2799	99.1	548	AA1992	Streptavidin/lucif
14	2799	99.1	548	AA1993	Mutant streptavidin
15	2785	98.7	788	AA1994	Antibody-Firefly 1
16	2701	95.7	552	AA1995	Chimeric firefly 1
17	2684	95.1	548	AA1996	Recombinant luciferase
18	2684	95.1	548	AA1997	Modified luciferase
19	2684	95.1	548	AA1998	Modified luciferase
20	2681	95.0	548	AA1999	Thermotable lucif
21	2680	94.9	548	AA2000	Thermotable lucif
22	2680	94.9	548	AA2001	Thermotable lucif
23	2678	94.9	548	AA2002	Thermotable lucif
24	2678	94.9	548	AA2003	Thermotable lucif
25	2678	94.9	548	AA2004	Thermotable lucif
26	2677	94.8	548	AA2005	Thermotable lucif
27	2676	94.8	548	AA2006	Thermotable lucif
28	2676	94.8	548	AA2007	Thermotable lucif
29	2674	94.7	548	AA2008	Thermotable lucif
30	2671	94.6	548	AA2009	Thermotable lucif
31	2671	94.6	548	AA2010	Thermotable lucif
32	2669	94.5	548	AA2011	Thermotable lucif
33	2590	91.7	552	AA2012	Thermotable lucif
34	2394	84.8	552	AA2013	Thermotable lucif
35	2343.5	83.0	548	AA2014	Thermotable lucif
36	1956.5	69.3	550	AA2015	Thermotable lucif
37	1956.5	69.3	550	AA2016	Thermotable lucif
38	1947.5	69.0	551	AA2017	Thermotable lucif
39	1946.5	68.9	550	AA2018	Thermotable lucif
40	1945.5	68.9	550	AA2019	Thermotable lucif
41	1945.5	68.9	550	AA2020	Thermotable lucif
42	1945.5	68.9	550	AA2021	Thermotable lucif
43	1942.5	68.8	550	AA2022	Thermotable lucif
44	1939.5	68.7	550	AA2023	Thermotable lucif
45	1939.5	68.7	550	AA2024	Thermotable lucif

ALIGNMENTS

RESULT 1
ID AA1980 standard; Protein: 548 AA.
AC AA1981;
DT 09-SEP-1999 (first entry)
DE L. lateralis luciferase protein sequence.
KW Luciferase; surfactant-tolerant; firefly; intracellular ATP assay.
OS Luciola lateralis.
PN WO9933997-A1.
PD 08-JUL-1999.
PF 24-DEC-1998; 98WO-JP05864.
PR 26-DEC-1997; 97JP-0361022.
PA (KIKK) KIKKOMAN CORP.
PI Hattori N, Murakami S;
DR WPI: 1999-419109/35.
PT N-PSDB; AAX84396.
PF Modified luciferase tolerant to surfactants and useful for assay of
PT Intracellular ATP
PS Claim 5; Page 43-48; 56pp; English.
XX

CC This sequence is a *Luciola lateralis* luciferase of the invention.
 CC The protein is a surfactant-tolerant luciferase which is derived from
 CC firefly luciferase by replacement of the glutamic acid residue at
 CC position 490 by another amino acid. The luciferase can be used in an
 CC assay of intracellular ATP in biological samples containing cells, for
 CC example for establishing the presence of cells, and their quantity, in
 CC the sample. The method can be applied to foodstuffs, drugs,
 CC agrochemicals, beverages and environmental samples. The activity of the
 CC surfactant-tolerant luciferase is not lowered by the presence of high
 CC concentrations of surfactant and therefore an assay of high accuracy can
 CC be carried out in a simple manner.

XX Sequence 548 AA:

Query Match 100.0%; Score 2823; DB 20; Length 548;
 Best Local Similarity 100.0%; Pred. No. 1e-235;
 Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MENMENDENIVGPEPPPIEESGAGOLRKYMDRYAKLGAIAFTNALTGVDYTAETLE 60
 DB 1 menmendenivgpeppfyleegsagaqlrkymdryaklgaiaftnalgvdytaeyle 60
 QY 61 KSCCLGEALKNYGLVVDGRALCSNCEEFIPVLAGLPIGVAVPPTNETYLRELIVHSL 120
 DB 61 kscclgealknyglvvdgralcsnceeffipvlaglfigvavptnetylrelivhsl 120
 QY 121 GISKPTIVFSSKKGIDKYITVQKTYVTAIKTIYILDSKVDYRGYQSDMNFIRKKNPQGFKG 180
 DB 121 giskptivfsskkgldkyltvqktyvtaiktiylilidskvdrygyqsdmnikknpqgfk 180
 QY 181 SSFKTVEVRKQOVALINSSGSTGLPKGYOLTHENIYTRFSHARPIYGNQVSGTAL 240
 DB 181 ssfktevnrkeqvalinssgstglpkgyoltheniytrfsharpiygnqvsgptail 240
 QY 241 TVPPEHGFEGMTTLGYLTCGFRIVMLTKFDETEFLKTLDYKCSSVILVPTLAILNRS 300
 DB 241 tvppfhgfgmttlgyltcgfrivmltkfdeetflktldykcassvllvptlailnrs 300
 QY 301 ELLDKYDLSNIVETASGAPLSKEIGEAVARFNPVGRQGYGLTETTSAILITPEGDGK 360
 DB 301 elldkydl snivelasgapslskeigeavarfnlpgvrgygltettsailitpegddk 360
 QY 361 PGASGKVVPLFKAKVIDLDTKKTGLPNRRGEVCVKGPMIMKYVDNPEATREITDEEGWL 420
 DB 361 pgasgkvvplfkakvidldtkktlgnrrgevcvkgpmlmkyvdnpeateitideegwl 420
 QY 421 HTGDIQYDEKHFPIYDRLSLKIKYGYOVPALESVLLQHPNIFDAGVAGVDPDIAG 480
 DB 421 htgdigydeekhfpiydrslkslikygyovpaellesvllqhpni fdagvagvdpdiag 480
 QY 481 ELPGAIVVLLKKGKSMTEKEMDYVASOVSNARKLRGVRVDEVKKGITGRI DKAIREI 540
 DB 481 elpgavvllkkgksmtekevmdyvasovsnarklrgvrvdevkkgitgrikdkairei 540
 QY 541 LKKFVAKM 548
 DB 541 lkkpvakm 548

RESULT 2

ID AAY22182 standard; Protein: 548 AA.

AC AAY22182;

DT 09-SEP-1999 (first entry)

DE L. lateralis luciferase protein sequence.

KW Luciferase; surfactant-tolerant; firefly; intracellular ATP assay.

QS Luciola lateralis.

XX PN W09393997-A1.
 XX PD 08-JUL-1999.
 XX PF 24-DEC-1998; 98WO-JP05864.
 XX PR 26-DEC-1997; 97JP-0361022.
 XX PA (KIKK) KIKKOMAN CORP.
 XX PI Hattori N, Murakami S;
 XX DR WPI; 1999-419109/35.
 XX DR N-PSDB; AAX84395.
 XX PT Modified luciferase tolerant to surfactants and useful for assay of
 XX PT intracellular ATP
 XX PS Claim 4; Page 34-39; 56pp; English.

CC This sequence is a *Luciola lateralis* luciferase of the invention.
 CC The protein is a surfactant-tolerant luciferase which is derived from
 CC firefly luciferase by replacement of the glutamic acid residue at
 CC position 490 by another amino acid. The luciferase can be used in an
 CC assay of intracellular ATP in biological samples containing cells, for
 CC example for establishing the presence of cells, and their quantity, in
 CC the sample. The method can be applied to foodstuffs, drugs,
 CC agrochemicals, beverages and environmental samples. The activity of the
 CC surfactant-tolerant luciferase is not lowered by the presence of high
 CC concentrations of surfactant and therefore an assay of high accuracy can
 CC be carried out in a simple manner.

SO Sequence 548 AA:

Query Match 99.9%; Score 2821; DB 20; Length 548;
 Best Local Similarity 99.8%; Pred. No. 1.5e-235;
 Matches 547; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MENMENDENIVGPEPPPIEESGAGOLRKYMDRYAKLGAIAFTNALTGVDYTAETLE 60
 DB 1 menmendenivgpeppfyleegsagaqlrkymdryaklgaiaftnalgvdytaeyle 60
 QY 61 KSCCLGEALKNYGLVVDGRALCSNCEEFIPVLAGLPIGVAVPPTNETYLRELIVHSL 120
 DB 61 kscclgealknyglvvdgralcsnceeffipvlaglfigvavptnetylrelivhsl 120
 QY 121 GISKPTIVFSSKKGIDKYITVQKTYVTAIKTIYILDSKVDYRGYQSDMNFIRKKNPQGFKG 180
 DB 121 giskptivfsskkgldkyltvqktyvtaiktiylilidskvdrygyqsdmnikknpqgfk 180
 QY 181 SSFKTVEVRKQOVALINSSGSTGLPKGYOLTHENIYTRFSHARPIYGNQVSGTAL 240
 DB 181 ssfktevnrkeqvalinssgstglpkgyoltheniytrfsharpiygnqvsgptail 240
 QY 241 TVPPEHGFEGMTTLGYLTCGFRIVMLTKFDETEFLKTLDYKCSSVILVPTLAILNRS 300
 DB 241 tvppfhgfgmttlgyltcgfrivmltkfdeetflktldykcassvllvptlailnrs 300
 QY 301 ELLDKYDLSNIVETASGAPLSKEIGEAVARFNPVGRQGYGLTETTSAILITPEGDGK 360
 DB 301 elldkydl snivelasgapslskeigeavarfnlpgvrgygltettsailitpegddk 360
 QY 361 PGASGKVVPLFKAKVIDLDTKKTGLPNRRGEVCVKGPMIMKYVDNPEATREITDEEGWL 420
 DB 361 pgasgkvvplfkakvidldtkktlgnrrgevcvkgpmlmkyvdnpeateitideegwl 420
 QY 421 HTGDIQYDEKHFPIYDRLSLKIKYGYOVPALESVLLQHPNIFDAGVAGVDPDIAG 480
 DB 421 htgdigydeekhfpiydrslkslikygyovpaellesvllqhpni fdagvagvdpdiag 480
 QY 481 ELPGAIVVLLKKGKSMTEKEMDYVASOVSNARKLRGVRVDEVKKGITGRI DKAIREI 540

Db 481 elpgavvvllkqksmtekevmdyvasvsnakrlirgvrivdevpkltgkidgkxalrei 540
 QY 541 LKKPVAKM 548
 Db 541 LKKPVAKM 548

RESULT 3

AAW04212
 ID AAW04212 standard; Protein; 548 AA.

AC AAW04212;

DT 12-NOV-1996 (first entry)

XX Luciola lateralis luciferase.

DE Mutant; streptavidin; luciferase; fusion protein;

KW Streptomyces avidinii; Luciola lateralis; firefly;

KM recombinant production; industry.

XX Luciola lateralis.

XX JP07289264-A.

XX 07-NOV-1995.

XX 27-APR-1994; 94JP-0090275.

XX 27-APR-1994; 94JP-0090275.

XX (KIKK) KIKKOMAN CORP.

XX WPI; 1996-015269/02.

XX New mutant streptavidin (SA) gene - fused to a firefly luciferase

PT gene, for the recombinant prepn. of a SA-FL fused protein

XX Disclosure; Page 10; 12pp; Japanese.

XX The present sequence is the L. lateralis luciferase, which can be

CC used in the construction of a streptavidin/luciferase fusion

CC protein, comprising a mutant streptomyces avidinii streptavidin

CC and the L. lateralis (firefly) luciferase. The fusion protein

CC can be prep'd. by inserting the recombinant DNA encoding it into a

CC Escherichia species microbe, culturing the transformed microbe in a

CC medium and collecting the fusion protein from the culture. The

CC fusion protein can be used in industry.

XX SQ Sequence 548 AA:

Query Match 99.8%; Score 2817; DB 17; Length 548;

Best Local Similarity 99.6%; Pred. No. 3.4e-235;

Matches 546; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MENMENDENIYGGPPPTPIEGSAGAO LRKYMRYAKGAIATNATLGVDTYAEYLE 60
 Db 1 menmendeniyygppptpiegsagaglrkymdyaklgatfnaltgydyaeyle 60
 QY 61 KSCCLGEALKNYGLVVDGRIALCSENCEFFIPVLAGEFIGVVAPTNEIYTLRELVHSL 120
 Db 61 kscclgealknyglvvdgrialcsenceeffipvlagelfigvvaptneiytlrelvhs1 120
 QY 121 GTSKPTIVSSKKGIDKRYTVOKTYTAIKTYIILDSKYDVRGYSMDNFIKKTPGFRKG 180
 Db 121 gtskptivsskkgidkrytvoktytaiktyiildskydyrgysmdnfikknpcgfrkg 180
 QY 181 SSPKTYEVNRKROVALIMNSGSGTGPKGVOLTHENIVTFESHADPIYGNQVSPGATL 240
 Db 181 ssftktyevnrkrovalimnsgsgtgpkgvoltthenivtfeshadpiygnqvspgtall 240

QY 241 TVPPEHHGFMETTLGYLTGCFRIVMTLKFDEETFLKTLQDYKCSSVILVPTLFAILNRS 300
 Db 241 tvpffhgmfttlyltcgfrivmtlktdeetflktldgkcssvllvptlfailnrs 300
 QY 301 ELDDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGICLTETTSALITTPBEDDK 360
 Db 301 elldkylslnlveiasgapslskeigearfnlpgvrgicltetsallitpbeddk 360
 QY 361 PGASGVVPLFEKAVYIDDTKKTGPNRGEVGCYKGMKMGVYDNNEATREIIDEEGWL 420
 Db 361 pgasgvvplfkavvldtktlgnrgevcykgpmlmkgyvdpneatreideegwl 420
 QY 421 HTGDIGYDEEKHEFTVDRLSLIKYKGYVPPAELESVLLQHPNIDAGVAGVPPIAG 480
 Db 421 htgdigydeekheftvdrslklykgyvppaelsevllqhpnidagvavpdpia 480
 QY 481 ELPGAUVVLLKKSMTKEKEMVDYVASQVSNKRLRGVRFVDEVPKLTGKIDGKALREI 540
 Db 481 elpgavvvllkksmtkekevmdyvasvsnakrlirgvrivdevpkltgkidgkxalrei 540
 QY 541 LKKPVAKM 548
 Db 541 LKKPVAKM 548

RESULT 4

AAW53884
 ID AAW53884 standard; Protein; 548 AA.

AC AAW53884;

DT 16-JUL-1998 (first entry)

XX Antibody-firefly luciferase fusion protein.

XX Firefly luciferase; antibody-luciferase fusion protein.

XX Synthetic.

XX Luciola cruciata.

XX JP09187281-A.

XX 22-JUL-1997.

XX 09-JAN-1996; 96JP-0001812.

XX 09-JAN-1996; 96JP-0001812.

XX (KIKK) KIKKOMAN CORP.

XX WPI; 1996-275089/25.

XX Antibody-firefly luciferase fused protein - and related products

PT i.e. firefly luciferase fused gene, recombinant DNA and its

PT preparation

XX Disclosure; Page 9-10; 17pp; Japanese.

XX This sequence is a fusion protein of the invention. The protein is

CC an antibody-firefly luciferase fusion protein, in which an antibody part

CC consisting of a peptide having antibody activity is combined with an

CC enzyme part consisting of firefly luciferase.

XX SQ Sequence 548 AA:

Query Match 99.8%; Score 2817; DB 19; Length 548;

Best Local Similarity 99.6%; Pred. No. 3.4e-235;

Matches 546; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MENMENDENIYGGPPPTPIEGSAGAO LRKYMRYAKGAIATNATLGVDTYAEYLE 60
 Db 1 menmendeniyygppptpiegsagaglrkymdyaklgatfnaltgydyaeyle 60

```

QY 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFPIVLAGLFTIGVAVAPNTNETLREIVHSL 120
DB 61 kscclgealknvglvdgrialcsenceeffipvlaglftigvavapntnetlreivhsl 120
QY 121 GISKPTIVSSSKGLDKVITVQKTVTAIKTIVILDSKVDYRGYOSMDNFKTKNTPQGFKG 180
DB 121 giskptivsskglkdvitvqktvtaiaktivildskvdrygysmdnfikntpgfkx 180
QY 181 SSFKTEVNRRKEQVALIMNSSGSTGLPKGVOLTHENIVTRFESHARDPIYGNOSPGTAL 240
DB 181 ssfktevnrkeqvalimnssgstglpkgyqlthenvltrfshardpiygnvsgptail 240
QY 241 TVPFRHGHGEMFTTGLYLCGFRIVMLTKFDEETPLKTLQDYKCSSVILPTLFAILNRS 300
DB 241 tvpfrhghgmfttglylfcgfrivmltkfdeetfllktdykcssvillptlfaillns 300
QY 301 ELLDKYDLNLYEIASGGAPLSKEIGEAVARRNLPGRVQSGVLTETTSATITTPGDDK 360
DB 301 elldkydlsllyeiasggaplskeigearrlnpgvrgyglteetsailltpgddk 360
QY 361 PGASGVVPLFRKAKVIDDITDKTKTLGPNRRGEVCVKGPMIMKGVNDPPEATREIIDEEGML 420
DB 361 pgasgvvplfrkavididitdktktlgpnrrgevcvkgpmlmkyvndppeatreideegwl 420
QY 421 HTGDIQYDDEKHFFIVDRLSLTKYKGYVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
DB 421 htgdigydeekhfifvdrkslkykgyvppaellesvllqhpnidagvagvdpdiag 480
QY 481 ELPGAIVVLLKKGSMTEKEVMDYASQVSNARKRLRGVRFVDEVPKGLTGKIDKAIKREI 540
DB 481 elpgavvvllkkgsmtekevmdyasqvsnakrlrgvrfvdevpkgltgkldgkaiarei 540
QY 541 LKKPVAKM 548
DB 541 lkkpvakm 548

RESULT 5
AAM9367
ID AAM9367 standard; Protein: 548 AA.
XX
AC AAM9367;
XX
DT 21-MAY-1999 (first entry)
XX
DE Mutant firefly luciferase protein #4.
XX
KW Bioluminescent protein; catalytic efficiency; stability; firefly;
KW luciferase; chimeric; luciola cruciata; luciola lateralis;
KW photinus pyralis; primer; PCR; amplification.
XX
OS Luciola lateralis.
XX
OS Synthetic.
XX
PN MO9902697-A1.
XX
PD 21-JAN-1999.
XX
PF 30-JUN-1998; 98MO-JP02936.
XX
PR 08-JUL-1997; 97US-0051917.
XX
PA (KIKK ) KIKKOMAN CORP.
XX
PI Hirokawa K, Kajiyama N, Murakami S;
XX
DR WPI: 1999-120898/10.
XX
DR N-PSDB; AAX25718.
XX
PT New bioluminescent protein with improved properties - has greater
PT catalytic efficiency and stability and is obtained by modification
PT of natural precursors

```

```

XX XX
PS Example 5; Page 36-39; 53pp; Japanese.
XX
CC The invention relates to the generation of bioluminescent proteins with
CC improved catalytic efficiency and stability. The proteins are generated
CC by addition, deletion or substitution of amino acids from a precursor
CC protein e.g. firefly luciferase, or by constructing chimeric luciferase
CC proteins from luciferase genes from e.g. luciola cruciata, L. lateralis
CC and photinus pyralis. This sequence represents an example of a mutated
CC luciferase protein of the invention.
XX
SQ Sequence 548 AA;

Query Match 99.7%; Score 2815; DB 20; Length 548;
Best Local Similarity 99.6%; Pred. No. 5e-235;
Matches 546; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MENMENDENIVYGPPEPPIREESAGAOQLRKYMDRAKIGATAFTALGVDVTVAEYLE 60
DB 1 menmendenvygppepfiyeesagaqlrkymdryakigataftalgvdytyaeyle 60
QY 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFPIVLAGLFTIGVAVAPNTNETLREIVHSL 120
DB 61 kscclgealknvglvdgrialcsenceeffipvlaglftigvavapntnetlreivhsl 120
QY 121 GISKPTIVSSSKGLDKVITVQKTVTAIKTIVILDSKVDYRGYOSMDNFKTKNTPQGFKG 180
DB 121 giskptivsskglkdvitvqktvtaiaktivildskvdrygysmdnfikntpgfkx 180
QY 181 SSFKTEVNRRKEQVALIMNSSGSTGLPKGVOLTHENIVTRFESHARDPIYGNOSPGTAL 240
DB 181 ssfktevnrkeqvalimnssgstglpkgyqlthenvltrfshardpiygnvsgptail 240
QY 241 TVPFRHGHGEMFTTGLYLCGFRIVMLTKFDEETPLKTLQDYKCSSVILPTLFAILNRS 300
DB 241 tvpfrhghgmfttglylfcgfrivmltkfdeetfllktdykcssvillptlfaillns 300
QY 301 ELLDKYDLNLYEIASGGAPLSKEIGEAVARRNLPGRVQSGVLTETTSATITTPGDDK 360
DB 301 elldkydlsllyeiasggaplskeigearrlnpgvrgyglteetsailltpgddk 360
QY 361 PGASGVVPLFRKAKVIDDITDKTKTLGPNRRGEVCVKGPMIMKGVNDPPEATREIIDEEGML 420
DB 361 pgasgvvplfrkavididitdktktlgpnrrgevcvkgpmlmkyvndppeatreideegwl 420
QY 421 HTGDIQYDDEKHFFIVDRLSLTKYKGYVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
DB 421 htgdigydeekhfifvdrkslkykgyvppaellesvllqhpnidagvagvdpdiag 480
QY 481 ELPGAIVVLLKKGSMTEKEVMDYASQVSNARKRLRGVRFVDEVPKGLTGKIDKAIKREI 540
DB 481 elpgavvvllkkgsmtekevmdyasqvsnakrlrgvrfvdevpkgltgkldgkaiarei 540
QY 541 LKKPVAKM 548
DB 541 lkkpvakm 548

RESULT 6
AAR30803
ID AAR30803 standard; Protein: 548 AA.
XX
AC AAR30803;
XX
DT 24-MAY-1993 (first entry)
XX
DE Thermostable luciferase (wild-type).
XX
KW Firefly; Luciola cruciata; GENJI; Luciola lateralis; HEIKE;
KW mutant; assay.
XX
OS Luciola lateralis.

```

```

XX Key Location/Qualifiers
FH Misc-difference 217
FT /note= "amino acid replaced by hydrophobic amino
FT acid, esp. Ile, Leu or Val, for heat stable
FT mutant".
XX
XX EP524448-A.
XX
XX 27-JAN-1993.
XX
XX 26-JUN-1992: 92EP-0110808.
XX
XX 27-JUN-1991: 91JP-0157117.
XX 29-NOV-1991: 91JP-0317064.
XX 22-MAY-1992: 92JP-0131057.
XX
XX (KIKK ) KIKKOMAN CORP.
XX
XX Elachi N, Naoki K;
XX
XX WPI: 1993-028553/04.
XX N-PSDB; AAQ34745.
XX
XX New DNA sequence of a thermostable luciferase - has the aminoacid
XX at position-217 of luciola cruciata or lateralis luciferase
XX replaced by a hydrophobic aminoacid
XX
XX Claim 10-12: Page 30 + 25-29; 33pp: English.
XX
XX A thermostable luciferase of a firefly, in which an amino acid
XX at the 217 position or an amino acid at the position equiv. to the
XX 217 position of the luciferase of Luciola cruciata (GENJI firefly)
XX of Luciola lateralis (HEIKE firefly) is replaced by a hydrophobic
XX amino acids, esp. Ile, Leu or Val, is claimed. The mutant luciferase
XX is identical in properties to the wild-type luciferase except that
XX it is stable when heated to high temps., e.g. 50 degrees C.
XX The luciferase can be used in assays for e.g. ATP.
XX
XX Sequence 548 AA:
XX
Query Match 99.7%; Score 2814; DB 14; Length 548;
Best Local Similarity 99.6%; Pred. No. 6.2e-235;
Matches 546; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 MENNENDENTIVYGPEPPYPIEGSAGAOIKRYMDRYAKGALFTNALTGVDYTAAYLE 60
DB 1 mennendentivygpeppyleegsagqlrkymdryaklgafalnltgvdyltaeyle 60
OY 61 KSCCLGEMAKWGLVVDGRALACSENCEEPFIVLAGLFTGAGVAPNNEYTLREIVHSL 120
DB 61 kscclgealmkwyglvvdgralacsenceefivlaglftgagvapnneyleivhsl 120
OY 121 GISKPTIYFSSKGLDKYITVQKTVTAIKTIVILDSVDYKGYOSMDNFKKNTPOGFKG 180
DB 121 gisrptiyfsskglkdyitvqktvtaiaktivildskvdrygysmhnfkknpggfk 180
OY 181 SSFKTEVNRKEQVALIMNSGSGLEPKGVOLTHEENTVTRSHARDPIYGNQVSPGAIL 240
DB 181 ssfktevnrkeqvalimnsgsgclpkgyvqlthenavtrshardpiygnvsgpail 240
OY 241 TVPFFHGFGEFTLGYTCGFRIVMLTKFDEETFLKTDYKSSVILVPTLFAINRS 300
DB 241 tvpffhgfgeftlgytcgfrivmltkfdeetflktldykssviltvptlfaainrs 300
OY 301 ELUDKYDLSNDELASGAPLSKETGEAVARRENLPVGRQGYGLTETTSALITPEGDDK 360
DB 301 elldkydlsnvelasgapsksetgeavarrenlpvgrqgygltettsaliltpedgdk 360
OY 361 PGASGRVPLRKAVYIDTCKTKLCPNRRGEVCVGPMLMGVYDNPATREIIDEEGWL 420
DB 361 pgasgrvplrkavyyidtktklcpnrrgevcvqpmllmgvydnpatreiideegwl 420

```

```

OY 421 HTGDIGYDEKHEFIYDRLSLIKRYKGYVPAELESVLLQHPNIDAGVAGPPPIAG 480
DB 421 htgdigydeekhefiydrlslikrykgyvpaelesvllqhpnidagvaggpppiag 480
OY 481 ELPCAVVVLKKGKSMTEKEVMDYVASOVSNARKLRGVRVEDEPKGLTGKIDGKALREI 540
DB 481 elpcavvvlkkgksmtekevmdyvasovsnarklrvgvrvedepkgltgkldgkalrei 540
OY 541 LKRPVAKM 548
DB 541 lkpvakm 548
RESULT 7
AAR33710
ID AAR33710 standard; Protein: 548 AA.
XX
XX AAR33710;
XX
XX 24-MAY-1993 (first entry)
XX
XX Thermostable luciferase (T217L,L,V).
XX
XX Firefly: Luciola cruciata; GENJI; Luciola lateralis; HEIKE;
XX mutant; assay.
XX
XX Luciola lateralis.
XX
XX Key Location/Qualifiers
FH Misc-difference 217
FT /label= ILE, LEU, VAL
FT
XX EP524448-A.
XX
XX 27-JAN-1993.
XX
XX 26-JUN-1992: 92EP-0110808.
XX
XX 27-JUN-1991: 91JP-0157117.
XX 29-NOV-1991: 91JP-0317064.
XX 22-MAY-1992: 92JP-0131057.
XX
XX (KIKK ) KIKKOMAN CORP.
XX
XX Elachi N, Naoki K;
XX
XX WPI: 1993-028553/04.
XX
XX New DNA sequence of a thermostable luciferase - has the aminoacid
XX at position-217 of luciola cruciata or lateralis luciferase
XX replaced by a hydrophobic aminoacid
XX
XX Claim 10-12: Page 30 + 25-29; 33pp: English.
XX
XX A thermostable luciferase of a firefly, in which an amino acid
XX at the 217 position or an amino acid at the position equiv. to the
XX 217 position of the luciferase of Luciola cruciata (GENJI firefly)
XX of Luciola lateralis (HEIKE firefly) is replaced by a hydrophobic
XX amino acids, esp. Ile, Leu or Val, is claimed. The mutant luciferase
XX is identical in properties to the wild-type luciferase except that
XX it is stable when heated to high temps., e.g. 50 degrees C.
XX The luciferase can be used in assays for e.g. ATP.
XX
XX Sequence 548 AA:
XX
Query Match 99.7%; Score 2814; DB 14; Length 548;
Best Local Similarity 99.6%; Pred. No. 6.2e-235;
Matches 546; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 MENNENDENTIVYGPEPPYPIEGSAGAOIKRYMDRYAKGALFTNALTGVDYTAAYLE 60

```

```

Db      1  menmendeniyygpepfyieegsagaqlrkymdryakigaiafinaltgvdtyaeyle 60
Qy      61  KSCCLGEALKNYGLVVDGRIALCSNCEEFPIVLAGLFVGAVAPRNEIYTLRELIVHSL 120
Db      61  KSCCLGEALKNYGLVVDGRIALCSNCEEFPIVLAGLFVGAVAPRNEIYTLRELIVHSL 120
Qy      121  GISKPTIVSSSKKGLDVKVITVOKTVAIKTIYILDSKVYRGYQSMDFEIKKNTPOGFKG 180
Db      121  GISKPTIVSSSKKGLDVKVITVOKTVAIKTIYILDSKVYRGYQSMDFEIKKNTPOGFKG 180
Qy      181  SSFKTEVNRKEQVALIMNSSGSTGLPKGVOLTHENIVTRFSHARPIYGNQVSPGTAIL 240
Db      181  SSFKTEVNRKEQVALIMNSSGSTGLPKGVOLTHENIVTRFSHARPIYGNQVSPGTAIL 240
Qy      241  TVPPIHGFMTTLGYITCGFRIVMLTKPDEETPLKTLQDYKCSSVILVPTLFAILNRS 300
Db      241  TVPPIHGFMTTLGYITCGFRIVMLTKPDEETPLKTLQDYKCSSVILVPTLFAILNRS 300
Qy      301  ELDDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGVTLETTSATITPBGDDK 360
Db      301  ELDDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGVTLETTSATITPBGDDK 360
Qy      361  PGASGVVPLFKAKVIDLDTKKTLLGNRRGEVCVKGPMIMKGYVNDPEATREIIEEGWL 420
Db      361  PGASGVVPLFKAKVIDLDTKKTLLGNRRGEVCVKGPMIMKGYVNDPEATREIIEEGWL 420
Qy      421  HTGDIGYDDEKHFFIYDRILSKLTKGYVPALESVLLQHPNIFDAGVAGVPPDIAG 480
Db      421  HTGDIGYDDEKHFFIYDRILSKLTKGYVPALESVLLQHPNIFDAGVAGVPPDIAG 480
Qy      481  ELPGAVVVLKKGKSMTEKEVMDYVASOVNAKRLRGVRFVDEVPKGLTGKIDGAKIREI 540
Db      481  ELPGAVVVLKKGKSMTEKEVMDYVASOVNAKRLRGVRFVDEVPKGLTGKIDGAKIREI 540
Qy      541  LKKPVAKM 548
Db      541  LKKPVAKM 548

```

RESULT 8

AAW12395
ID AAW12395 standard; Protein: 636 AA.

AC AAW12395;
DT 20-MAY-1997 (first entry)

DE Firefly luciferase conjugated with 87 aa biotin peptide at 3' end.
KW Fusion protein; firefly; lucifera; luciferase; biotinylation;
KW wild type; E.coli; bioluminescence assay; quantification; ligand;
KW receptor.

OS Synthetic.

XX Key
XX Location/Qualifiers

FT Region
FT 1..547
FT /note= "Firefly luciferase portion of fusion protein"

FT Region
FT 548..549
FT /note= "encoded by linker sequence used to construct fusion gene"

FT Region
FT 549..636
FT /note= "biotin peptide portion of fusion protein"

XX JP08308578-A.

XX 26-NOV-1996.

XX 24-APR-1995; 95JP-0098857.

XX 14-MAR-1995; 95JP-0054625.

XX 27-JUL-1994; 94JP-0193798.

PA (KIKK) KIKKOMAN CORP.
XX WPI: 1997-059697/06.
DR N-PSDB: AAT63269.
XX
PT Fusion protein comprising firefly luciferase and biotinylated
PT peptide - useful in a bio-luminescent analytical method for
XX quantifying ligands
XX
PS Claim 2; Page 12-13; 13pp; Japanese.
XX
CC This is the amino acid sequence of a novel fusion protein which comprises
CC the firefly (Luciola lateralis) luciferase protein and an 87 amino acid
CC biotinylated peptide designated the E.coli biotin carboxy-carrier protein
CC (BCCP-87). The fusion gene was generated by firstly amplifying the
CC sequence encoding the E.coli BCCP-87 from the E.coli genome and inserting
CC the resultant sequence into the 3' end of the firefly luciferase gene in
CC plasmid pHLf230. The biotin peptide sequence replaces the C-terminal Met
CC residue of the wild type luciferase sequence. The plasmid was then
CC introduced into E.coli JM101 for production of the fusion protein. The
CC novel protein can be used in bioluminescence assays to quantify
CC luciferase ligands which may modulate binding of luciferase to its
CC receptor.
CC
SQ Sequence 636 AA;

Query Match 99.6%; Score 2812; DB 18; Length 636;
Best Local Similarity 99.6%; Pred. No. 1,1e-234;
Matches 545; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1  MENMENDENIYYGPEPFYIEEGSAGAO LRKYMDRYAKIGAIAFINALTGVDYTYAEYLE 60
Db      1  menmendeniyygpepfyieegsagaqlrkymdryakigaiafinaltgvdtyaeyle 60
Qy      61  KSCCLGEALKNYGLVVDGRIALCSNCEEFPIVLAGLFVGAVAPRNEIYTLRELIVHSL 120
Db      61  KSCCLGEALKNYGLVVDGRIALCSNCEEFPIVLAGLFVGAVAPRNEIYTLRELIVHSL 120
Qy      121  GISKPTIVSSSKKGLDVKVITVOKTVAIKTIYILDSKVYRGYQSMDFEIKKNTPOGFKG 180
Db      121  GISKPTIVSSSKKGLDVKVITVOKTVAIKTIYILDSKVYRGYQSMDFEIKKNTPOGFKG 180
Qy      181  SSFKTEVNRKEQVALIMNSSGSTGLPKGVOLTHENIVTRFSHARPIYGNQVSPGTAIL 240
Db      181  SSFKTEVNRKEQVALIMNSSGSTGLPKGVOLTHENIVTRFSHARPIYGNQVSPGTAIL 240
Qy      241  TVPPIHGFMTTLGYITCGFRIVMLTKPDEETPLKTLQDYKCSSVILVPTLFAILNRS 300
Db      241  TVPPIHGFMTTLGYITCGFRIVMLTKPDEETPLKTLQDYKCSSVILVPTLFAILNRS 300
Qy      301  ELDDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGVTLETTSATITPBGDDK 360
Db      301  ELDDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGVTLETTSATITPBGDDK 360
Qy      361  PGASGVVPLFKAKVIDLDTKKTLLGNRRGEVCVKGPMIMKGYVNDPEATREIIEEGWL 420
Db      361  PGASGVVPLFKAKVIDLDTKKTLLGNRRGEVCVKGPMIMKGYVNDPEATREIIEEGWL 420
Qy      421  HTGDIGYDDEKHFFIYDRILSKLTKGYVPALESVLLQHPNIFDAGVAGVPPDIAG 480
Db      421  HTGDIGYDDEKHFFIYDRILSKLTKGYVPALESVLLQHPNIFDAGVAGVPPDIAG 480
Qy      481  ELPGAVVVLKKGKSMTEKEVMDYVASOVNAKRLRGVRFVDEVPKGLTGKIDGAKIREI 540
Db      481  ELPGAVVVLKKGKSMTEKEVMDYVASOVNAKRLRGVRFVDEVPKGLTGKIDGAKIREI 540
Qy      541  LKKPVAK 547
Db      541  LKKPVAK 547

```

RESULT 9

AAW48855
ID AAW48855 standard; Protein; 640 AA.
XX
AC AAW48855;
XX
DT *10-SEP-1998 (first entry)
XX
DE Luciferase-lyso-staphin fused protein.
XX
KW Luciferase-lyso-staphin fused protein; chimeric; bioluminescence.
XX
OS Chimeric - Staphylococcus simulans.
XX
OS Chimeric - Luciola lateralis.
XX
FH Key Location/Qualifiers
FT Domain 1..549
FT /note= "Luciferase protein"
FT 550..640
FT Domain /note= "lyso-staphin protein"
XX
PN JP10150991-A.
PD 09-JUN-1998.
XX
PE 25-NOV-1996; 96JP-0328042.
XX
PR 25-NOV-1996; 96JP-0328042.
XX
PA (KIKK) KIKKOMAN CORP.
XX
DR WPI: 1998-379994/33.
DR N-PSDB: AAV32467.
XX
PT New luciferase-lyso-staphin fused protein - useful for
PT bioluminescence analysis
XX
PS Claim 4; Pages 7-9; 10pp; Japanese.
XX
CC The invention claims for a luciferase-lyso-staphin fused gene which
CC encodes a fused protein in which a peptide part consisting of
CC luciferase, from Luciola lateralis, is connected to a peptide part
CC consisting of lyso-staphin from Staphylococcus simulans. The method
CC described by the invention can be used to prepare a luciferase-
CC lyso-staphin fused protein efficiently. The fused protein is useful
CC for bioluminescence analysis.
XX
SQ Sequence 640 AA;
XX
Query Match 99.6%; Score 2812; DB 19; Length 640;
Best Local Similarity 99.6%; Pred. No. 1.2e-234;
Matches 545; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 MENMENDENIYVGPPEFYPIEGSAGAO LRKYMDRYAKGALFTNALTGVDYTAEYLE 60
DB 1 memmendeniygpepfypleegsagaqlrkymdryaklgafalnalgdytaeyle 60
OY 61 KSCCGEALKNYGLVNDGRALCSNCEEFRTYVLAGLFTGVGAFTNTEITLREIVHSL 120
DB 61 kscclgealknyglvndgralcsnceeffrtvylaglfvgvgaftnteytlreivhsl 120
OY 121 GISKPTIVFSSKRGDKVITVOKTATKTIIVILDSKVDYGYOSMONFIKKNTPQGRFG 180
DB 121 giskptivfsskrgdkvityvoktatkitiivildskvdgygysmonfikkntppqgrfg 180
OY 181 SSFKTEVNRREQVALIMNSSGSTGLPKGVOLTHEINIVTRESHARDPIYGNQVSGTAL 240
DB 181 ssfktevnrrkeqvalimnssgstglpkgvoltneinivtreshardpiygnqvsgtal 240
OY 241 TVVPPHNGFGMTTIGYLTGCFRIYMLTKRFDENFTLKLQDYKCSSVILVPLFAILNRS 300
DB 241 tvvpfhngfgmttigyltgcfritymltkrfdenftlklqdykcssvilvplfai lnrs 300
OY 300

OY 301 ELLDKYDLSNLVEIASGAPLSKEIGEA VARRNLPVGRQGYGLTETTSAIIITPEGDK 360
DB 301 elldkydl snlveiasgapslskeigearrnlpvgrqgyglte tsaaiitpegdk 360
OY 361 PGASGRVPLFRKAVYIDLDTKKTLGPNRRGEVCKGPM LMGYVNDPEATREIIDEGWL 420
DB 361 pgasgrvplfrkav yidldtkktlgpnrrgevcvkgpmlmkyvndpeatreideegwl 420
OY 421 HMGDITGYDEEKHFFIVDYLKSLIKYKGYQVPAPAELESVLLOHPNIFDAGVAGVDPDPIAG 480
DB 421 hmgditydeekhffivdylkslikygyvppael esvlqhpnifdagvavdpdpiag 480
OY 481 ELPGAVVVLKKGSKWTEKEMDVASQVSNARKRLRGVAFVDEVPKGLTGKIDGKAIREI 540
DB 481 elpgavvvllkkg skwtekevmdivasqvsnarkrlrgvafvdevpkgltgk idgkairei 540
OY 541 LKKPVAK 547
DB 541 lkkpvak 547
RESULT 10
AAR03731
ID AAR03731 standard; protein; 548 AA.
XX
AC AAR03731;
XX
DT 22-JUL-1990 (first entry)
XX
DE Luciola lateralis recombinant luciferase.
XX
KW Luciferase; enzyme; ATP; plasmid pHLf7; firefly.
XX
OS Luciola lateralis.
XX
PN EP353464-A.
XX
PD 07-FEB-1990.
XX
PE 30-JUN-1989; 89EP-0111958.
XX
PR 22-DEC-1988; 88JP-0162402.
PR 01-JUL-1988; 88JP-0162402.
XX
PA (KIKK) KIKKOMAN CORP.
XX
PI Tatsumi H, Kajiyama N, Nakano E;
XX
DR WPI: 1990-038240/06.
DR N-PSDB: AAC03257.
XX
PT New gene encoding luciferase -
PT derived from Luciola lateralis and expressed in and purified
PT from Escherichia coli genus cells.
XX
PS Claim 2; page 21; 43pp; English.
XX
CC This enzyme is useful for determining ATP levels. It is
CC contained in plasmid pHLf7.
XX
SQ Sequence 548 AA;
XX
Query Match 99.6%; Score 2811; DB 11; Length 548;
Best Local Similarity 99.5%; Pred. No. 1.1e-234;
Matches 545; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 1 MENMENDENIYVGPPEFYPIEGSAGAO LRKYMDRYAKGALFTNALTGVDYTAEYLE 60
DB 1 memmendeniygpepfypleegsagaqlrkymdryaklgafalnalgdytaeyle 60
OY 61 KSCCGEALKNYGLVNDGRALCSNCEEFRTYVLAGLFTGVGAFTNTEITLREIVHSL 120
DB 61 kscclgealknyglvndgralcsnceeffrtvylaglfvgvgaftnteytlreivhsl 120
DB 120

```

Db      61 kscclgealknvglvvdgrialcsenceeffivlaglfigvavaptneljlrelvhs1 120
QY      121 GISKPTIVFSSKKGLDKVITVQKTVTAIKTIVILDSKVDYRGVQSMDFIKKMTPOGFKG 180
      122 glskprlvtsskkgldkvlvtqktvtaiktivildskvdrygysmdnflkntlpqfkg 180
QY      181 SSKRYEVNKRKEVALIMNSSGSTGLPKGVQLTHENIVTRESHARDPIYGNQVSPGTAIL 240
      181 sskfkevntrkeqvalilnmssgstglpkgyqlthenavtrfshardpdygnvsgptail 240
QY      241 IVVPEFHGFMFTLLGYLTCGFRIVMLTKRDEFTFLTKTLDYKCSSVILVPTLFAILNRS 300
      241 tvvfhngfmgftllgyllcgfrivmltkfdeetflklqdykcssvllvptlfaillns 300
QY      301 ELIDKYDLSNLVEIASGAPLSKEIGEAARFNLPGVRQGYLTETTSAILITPEGDDK 360
      301 elldkydlslnlvelaasgapslskeigeavarfnlpgvrqgylltettssailitpegddk 360
QY      361 PGASGVVPLFKAKVITDTRKTLGPNRGEVGVKGPMLMKGVNDPEATREIIDEGBWL 420
      361 pgasgvvplfkakvldtrktlgpnrrgevcvkgpmlmkgyvndpeatreiideegwll 420
QY      421 HFGDIDGYDEEKNHFTVDRLSLIRKGYOVPAELSVLLOHNPIDAGVAGVPPPIAG 480
      421 hfgdidgydeeknhftvdrslslikkygyvppaelsevllohnplfdagvagvpppiag 480
QY      481 ELPGAVVVLKKGKSMTEKEVMDYVASOVSNARKLRGVRFEVDEPKLTGKIDGKAIREI 540
      481 elpgavvvllkkgksmtekevmdyvasovsnarklrgrvrfdevpkltgkldgkairrel 540
QY      541 LKKPVAKM 548
      541 lkkpvakm 548
Db

```

RESULT 11
AAM12394
ID AAM12394 standard; Protein: 568 AA.
AC AAM12394;
XX
DT 20-MAY-1997 (first entry)
XX
DE Firefly luciferase conjugated with 23 aa biotin peptide at N-terminus.
XX
KM Fusion protein; firefly; luciola lateralis; luciferase; biotinylation;
KM wild type; E.coli; bioluminescence assay; quantification; ligand;
KM receptor.
XX
OS Synthetic.
XX
FH Key
FH Region
FT
FT Region
FT
FT Region
XX
XX JP08308578-A.
XX
XX 26-NOV-1996.
XX
XX 24-APR-1995; 95JP-0098857.
XX
XX 14-MAR-1995; 95JP-0054625.
XX
XX 27-JUL-1994; 94JP-0193798.
XX
XX (KIKK) KIKKOMAN CORP.
XX
XX WPI: 1997-059697/06.
XX
XX N-PSDB: AAT63268.
XX
XX Fusion protein comprising firefly luciferase and biotinylated
XX peptide - useful in a bio-luminescent analytical method for
PT

```

PT      quantifying ligands
XX
XX Claim 2; Page 10-11; 13pp; Japanese.
XX
XX This is the amino acid sequence of a novel fusion protein comprising
CC the firefly (Luciola lateralis) luciferase protein and a 23 amino acid
CC biotinylated peptide designated biotin peptide #84. The gene was
CC generated by firstly annealing the sequence encoding peptide #84
CC (AAT63270) and its complementary sequence, then inserting the resultant
CC double stranded sequence into the 5' end of the luciferase gene in
CC plasmid pHL108. The biotin peptide sequence replaces the first 4 amino
CC acids of the wild type luciferase sequence. The plasmid was then
CC introduced into E.coli JM101 for production of the fusion protein. The
CC novel protein can be used in bioluminescence assays to quantify
CC luciferase ligands which may modulate binding of luciferase to its
CC receptor.
XX
SQ      Sequence 568 AA:

```

Query Match 99.1%; Score 2799; DB 18; Length 568;
Best Local Similarity 99.3%; Pred. No. 1.3e-233;
Matches 542; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

QY      3 NMENDENIYGPPEPPIYBEGSAGAO LRKYMRYAKGAIATFNALTGVDYTAEYLEKS 62
      3 slendeniivgpepfybiegsagaglrkymdryaklgatfnaltgvdtyaeyleks 62
Db
QY      63 CCLGEALKMYGLVNDRIALCSNCEEFIPVLAGIFIGVAVAPTNELJLRELVHSLGI 122
      63 cclgealkmyglvndrialcsenceeffivlaglfigvavaptneljlrelvhs1g1 122
Db
QY      83 kscclgealknvglvvdgrialcsenceeffivlaglfigvavaptneljlrelvhs1g1 142
      83 kscclgealknvglvvdgrialcsenceeffivlaglfigvavaptneljlrelvhs1g1 142
QY      123 SKPTIVFSSKKGLDKVITVQKTVTAIKTIVILDSKVDYRGVQSMDFIKKMTPOGFKGS 182
      123 skptivfsskkgldkvlvtqktvtaiktivildskvdrygysmdnflkntlpqfkgss 182
Db
QY      143 SSKRYEVNKRKEVALIMNSSGSTGLPKGVQLTHENIVTRESHARDPIYGNQVSPGTAIL 242
      143 sskryevnkrkevalimnssgstglpkgvqlthenivtreshardpdygnvsgptailv 242
Db
QY      181 SSKRYEVNKRKEVALIMNSSGSTGLPKGVQLTHENIVTRESHARDPIYGNQVSPGTAIL 242
      181 sskfkevntrkeqvalilnmssgstglpkgyqlthenavtrfshardpdygnvsgptail 242
Db
QY      203 FIVEVNRKEQVALIMNSSGSTGLPKGVQLTHENIVTRESHARDPIYGNQVSPGTAIL 262
      203 fkevntrkeqvalilnmssgstglpkgyqlthenavtrfshardpdygnvsgptailv 262
QY      243 VPFHNGFEGMFTLLGYLTCGFRIVMLTKRDEFTFLTKTLDYKCSSVILVPTLFAILNRS 302
      243 vpfhngfmgftllgyllcgfrivmltkfdeetflklqdykcssvllvptlfaillnrsel 302
Db
QY      263 VPFHNGFEGMFTLLGYLTCGFRIVMLTKRDEFTFLTKTLDYKCSSVILVPTLFAILNRS 322
      263 vpfhngfmgftllgyllcgfrivmltkfdeetflklqdykcssvllvptlfaillnrsel 322
QY      303 ELIDKYDLSNLVEIASGAPLSKEIGEAARFNLPGVRQGYLTETTSAILITPEGDDK 362
      303 elldkydlslnlvelaasgapslskeigeavarfnlpgvrqgylltettssailitpegddk 362
Db
QY      323 LKYLDSLNLVEIASGAPLSKEIGEAARFNLPGVRQGYLTETTSAILITPEGDDK 382
      323 lkyldslnlvelaasgapslskeigeavarfnlpgvrqgylltettssailitpegddk 382
QY      363 ASGRVYVPLFKAKVITDTRKTLGPNRGEVGVKGPMLMKGVNDPEATREIIDEGBWL 422
      363 asgrvyvplfkakvldtrktlgpnrrgevcvkgpmlmkgyvndpeatreiideegwllht 422
Db
QY      383 ASGRVYVPLFKAKVITDTRKTLGPNRGEVGVKGPMLMKGVNDPEATREIIDEGBWL 442
      383 asgrvyvplfkakvldtrktlgpnrrgevcvkgpmlmkgyvndpeatreiideegwllht 442
QY      423 GDIGYVDEEKNHFTVDRLSLIRKGYOVPAELSVLLOHNPIDAGVAGVPPPIAG 482
      423 gdigyvdeeknhftvdrslslikkygyvppaelsevllohnplfdagvagvpppiagel 482
Db
QY      443 GDIGYVDEEKNHFTVDRLSLIRKGYOVPAELSVLLOHNPIDAGVAGVPPPIAG 502
      443 gdigyvdeeknhftvdrslslikkygyvppaelsevllohnplfdagvagvpppiagel 502
QY      483 PGAVVVLKKGKSMTEKEVMDYVASOVSNARKLRGVRFEVDEPKLTGKIDGKAIREI 542
      483 pgavvvllkkgksmtekevmdyvasovsnarklrgrvrfdevpkltgkldgkairrel 542
Db
QY      503 PGAVVVLKKGKSMTEKEVMDYVASOVSNARKLRGVRFEVDEPKLTGKIDGKAIREI 562
      503 pgavvvllkkgksmtekevmdyvasovsnarklrgrvrfdevpkltgkldgkairrel 562
QY      543 KPVAKM 548
      543 kpvakm 548
Db

```

RESULT 12
AAM53883
ID AAM53883 standard; Protein: 568 AA.
AC AAM53883;
XX
XX
XX 16-JUL-1998 (first entry)
XX
XX
XX

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2001, 01:51:50 ; Search time 36.66 Seconds

(without alignments)
1138.670 Million cell updates/sec

Title: US-09-581-241-4

Perfect score: 2823
Sequence: 1 MENMENDENIVYGPPEFYPI.....TGKIDKAIKEIKKKVAKM 548

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2814	99.7	548	1 S23437	Photinus-luciferin
2	2684	95.1	548	1 S20181	Photinus-luciferin
3	2335.5	82.7	548	2 S33788	Photinus-luciferin
4	1965.5	69.6	547	2 S62787	Photinus-luciferin
5	1945.5	68.9	550	1 A26772	Photinus-luciferin
6	1365	48.4	543	2 S29354	Photinus-luciferin
7	1365	48.4	543	2 S29354	Photinus-luciferin
8	1361	48.2	543	2 S29355	Photinus-luciferin
9	1354	48.0	543	2 S29352	Photinus-luciferin
10	768.5	27.2	544	1 S01667	4-coumarate--CoA 1
11	760.5	26.9	544	2 S15695	4-coumarate--CoA 1
12	758	26.9	557	2 T07909	4-coumarate--CoA 1
13	755.5	26.8	548	2 T07908	4-coumarate--CoA 1
14	755	26.7	544	2 H85064	4-coumarate--CoA 1
15	748.5	26.5	542	2 T03789	4-coumarate--CoA 1
16	740.5	26.2	542	1 T02074	4-coumarate--CoA 1
17	734.5	26.0	535	2 T08074	4-coumarate--CoA 1
18	734.5	26.0	545	1 A39827	4-coumarate--CoA 1
19	733	26.0	541	2 D96674	hypothetical prote
20	732	25.9	563	1 J00311	4-coumarate--CoA 1
21	727.5	25.8	545	1 B39827	4-coumarate--CoA 1
22	721	25.5	1549	2 D85338	protein F5M15.18
23	720.5	25.5	561	2 S57784	4-coumarate--CoA 1
24	720	25.5	544	2 T20741	hypothetical prote
25	715	25.3	570	2 T08075	4-coumarate--CoA 1
26	689.5	24.4	569	2 T03390	4-coumarate--CoA 1
27	682	24.2	537	2 T09755	4-coumarate--CoA 1
28	676.5	24.0	542	2 B96654	hypothetical prote
29	674	23.9	537	2 T09710	4-coumarate--CoA 1

30	661.5	23.4	423	2 PQ0772	4-coumarate--CoA 1
31	635	22.5	552	2 E69438	probable fatty-aci
32	634.5	22.5	566	2 F85214	4-coumarate--CoA 1
33	634.5	22.5	566	2 T05038	4-coumarate--CoA 1
34	602	21.3	566	2 H84037	long-chain fatty-a
35	599.5	21.2	569	2 C69471	probable fatty-aci
36	596.5	21.1	564	2 G84037	long-chain fatty-a
37	587.5	20.8	293	2 S31705	4-coumarate--CoA 1
38	579.5	20.5	593	2 E69378	probable acid--CoA
39	575	20.4	548	2 D69187	probable acid--CoA
40	573	20.3	540	2 T18841	hypothetical prote
41	572.5	20.3	560	2 D69649	hypothetical prote
42	569	20.2	561	1 S41589	long-chain-fatty-a
43	569	20.2	561	2 F85791	hypothetical prote
44	566.5	20.1	584	2 C75364	probable long-chain
45	563	19.9	546	2 C83791	acid-CoA ligase BH

ALIGNMENTS

RESULT 1

S23437 Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) - Japanese firefly

N:Alternate names: firefly luciferase

C:Species: Luciola lateralis (Japanese firefly)

C>Date: 22-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 02-Jun-2000

C:Accession: S23437; S57417

R:Tatsumi, H.; Kajiyama, N.; Nakano, E.

Biochim. Biophys. Acta 1131, 161-165, 1992

A:Title: Molecular cloning and expression in Escherichia coli of a cDNA clone encodin

A:Accession: S23437; M01D:92305054

A:Residues: 1-548 <TAT>

A:Cross-references: EMBL:X66919; NID:99526; PIDN:CAA47358.1; PID:99527

R:Cho, K.; Choi, Y.; Boo, K.

submitted to the EMBL Data Library, June 1995

A:Description: Molecular cloning of gene for luciferase in Luciola lateralis.

A:Reference number: S57417

A:Accession: S57417

A:Residue type: DNA

A:Residues: 1-4, 'D', '6-14, 'K', '16-145, 'N', '147-175, 'P', '177-505, 'G', '507-548 <CHO>

A:Cross-references: EMBL:Z49891; NID:9871400; PIDN:CAA90072.1; PID:9871401

C:Genetics:

C:Introns: 44/3; 114/1; 225/2; 332/2; 450/3; 506/1

C:Function:

A:Description: catalyzes the irreversible reaction of luciferin with oxygen and ATP to

C:Note: magnesium is required; color of emitted light varies among species from green

C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology

C:Keywords: ATP; luminescence; magnesium; monooxygenase; oxidoreductase; peroxisome

F:73-537/Domain: acetate--CoA ligase homology <ACLD>

Query Match 99.7%; Score 2814; DB 1; Length 548;

Best local similarity 99.6%; Pred. No. 1.2e-192;

Matches 546; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

DB 1 MENMENDENIVYGPPEFYPIEBGSAQOLRYMDRYAKGAIATNALTGVDYTYAYLE 60

DB 1 KSCCGEALKNYGLVVDGRIMLCSENCEFFIPVAGFVGVAAPTNEITLRELVS 120

DB 61 KSCCGEALKNYGLVVDGRIMLCSENCEFFIPVAGFVGVAAPTNEITLRELVS 120

DB 121 GISKPTIVFSSKKGDKVTYVQKTYTAIKTYIILDSKYDYGOSMDNFIRKMPQCKG 180

DB 121 GISKPTIVFSSKKGDKVTYVQKTYTAIKTYIILDSKYDYGOSMDNFIRKMPQCKG 180

DB 181 SSFKTVEVNRKQVALINSSGSLPKGVOLTHENLVTRSHADPTLYGNVSPGTA 240

DB 181 SSFKTVEVNRKQVALINSSGSLPKGVOLTHENLVTRSHADPTLYGNVSPGTA 240

QY 241 TVVPHHGFMTTGLYTCGFRIVMLTKPDEETFLKTLQDYKCSSVILVPTLAILNRS 300
 |||||
 Db 241 TVVPHHGFMTTGLYTCGFRIVMLTKPDEETFLKTLQDYKCSSVILVPTLAILNRS 300
 QY 301 ELLDKYLSNLVEIASGAPLSKEIGAVARFNPGLRGVGLTETTSAILITPEGDGK 360
 |||||
 Db 301 ELLDKYLSNLVEIASGAPLSKEIGAVARFNPGLRGVGLTETTSAILITPEGDGK 360
 QY 361 PGASGVVPLFKAKVIDLDTKTLGPNRRGECVCKGPMKGVNDPEARETIIDEEGWL 420
 |||||
 Db 361 PGASGVVPLFKAKVIDLDTKTLGPNRRGECVCKGPMKGVNDPEARETIIDEEGWL 420
 QY 421 HTGDIGYDEEKHFIVDRLSLTKYQVPPAELESVLLQHPNIFDAGVAGVPPPIAG 480
 |||||
 Db 421 HTGDIGYDEEKHFIVDRLSLTKYQVPPAELESVLLQHPNIFDAGVAGVPPPIAG 480
 QY 481 ELPGAVVVLKKGSKMTEKEVMQVYASQVSNAKRLRGVRFVDEVPKGLTGKIDGKAIREI 540
 |||||
 Db 481 ELPGAVVVLKKGSKMTEKEVMQVYASQVSNAKRLRGVRFVDEVPKGLTGKIDGKAIREI 540
 QY 541 LKRPVAKM 548
 |||||
 Db 541 LKRPVAKM 548

RESULT 2

Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) - Genj1 firefly
 J50181
 N:Alternate names: firefly luciferase
 C:Species: Luciola cruciata (Genj1 firefly)
 C>Date: 07-Jun-1990 #sequence_revision 20-Aug-1994 #text_change 26-May-2000
 C:Accession: J50181
 R:Masuda, T.; Tatsumi, H.; Nakano, E.
 Gene 77, 265-270, 1989
 A:Title: Cloning and sequence analysis of cDNA for luciferase of a Japanese firefly, Luc
 A:Reference number: J50181; MUID:89326143
 A:Accession: J50181
 A:Molecule type: mRNA
 A:Residues: 1-548 <MAS>
 A:Cross-references: GB:M26194; NID:9159050; PIDN:AAA29135.1; PID:9159051
 C:Comment: This protein catalyzes the oxidation of luciferin in the presence of ATP, oxy
 species from green to yellow.
 C:Superfamily: 4-comumate--CoA ligase; acetate--CoA ligase homology
 C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase
 F:73-537/Domain: acetate--CoA ligase homology <ACL>

Query Match 95.1%; Score 2684; DB 1; Length 548;
 Best Local Similarity 93.4%; Pred. No. 2,1e-183;

Matches 512; Conservative 26; Mismatches 10; Indels 0; Gaps 0;

QY 1 MENMENDENIVYGPPEPPYPIEESGAGAOILRKYMDRYAKGALAFNALGVDTYAEYLE 60
 |||||
 Db 1 MENMENDENIVYGPPEPPYPIEESGAGAOILRKYMDRYAKGALAFNALGVDTYAEYLE 60
 QY 61 KSCCGEALKNGLVVDGRIALCSENCEFFIPVLAGLFGVAPVNEIYTLRELHSL 120
 |||||
 Db 61 KSCCGEALKNGLVVDGRIALCSENCEFFIPVLAGLFGVAPVNEIYTLRELHSL 120
 QY 121 GISKPTIVFSSKKGLDKVITVQKTAIKTIYIIDSQKDYRGYQSDNFIKKNPQGGFK 180
 |||||
 Db 121 GISKPTIVFSSKKGLDKVITVQKTAIKTIYIIDSQKDYRGYQSDNFIKKNPQGGFK 180
 QY 181 SSFKIVEVNRKEQVALINSSSGSTGLPKGYOLTHEMLVTRSHADPIYGNQVSGTAVL 240
 |||||
 Db 181 SSFKIVEVNRKEQVALINSSSGSTGLPKGYOLTHEMLVTRSHADPIYGNQVSGTAVL 240
 QY 241 TVVPHHGFMTTGLYTCGFRIVMLTKPDEETFLKTLQDYKCSSVILVPTLAILNRS 300
 |||||
 Db 241 TVVPHHGFMTTGLYTCGFRIVMLTKPDEETFLKTLQDYKCSSVILVPTLAILNRS 300
 QY 301 ELLDKYLSNLVEIASGAPLSKEIGAVARFNPGLRGVGLTETTSAILITPEGDGK 360
 |||||

|||
 Db 301 ELLDKYLSNLVEIASGAPLSKEIGAVARFNPGLRGVGLTETTSAILITPEGDGK 360
 QY 361 PGASGVVPLFKAKVIDLDTKTLGPNRRGECVCKGPMKGVNDPEARETIIDEEGWL 420
 |||||
 Db 361 PGASGVVPLFKAKVIDLDTKTLGPNRRGECVCKGPMKGVNDPEARETIIDEEGWL 420
 QY 421 HTGDIGYDEEKHFIVDRLSLTKYQVPPAELESVLLQHPNIFDAGVAGVPPPIAG 480
 |||||
 Db 421 HTGDIGYDEEKHFIVDRLSLTKYQVPPAELESVLLQHPNIFDAGVAGVPPPIAG 480
 QY 481 ELPGAVVVLKKGSKMTEKEVMQVYASQVSNAKRLRGVRFVDEVPKGLTGKIDGKAIREI 540
 |||||
 Db 481 ELPGAVVVLKKGSKMTEKEVMQVYASQVSNAKRLRGVRFVDEVPKGLTGKIDGKAIREI 540
 QY 541 LKRPVAKM 548
 |||||
 Db 541 LKRPVAKM 548

RESULT 3

Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) [validated] - sou
 S33788
 N:Alternate names: firefly luciferase
 C:Species: Luciola mingrellica (southern Russian firefly)
 C>Date: 02-Dec-1993 #sequence_revision 03-Nov-1995 #text_change 02-Jun-2000
 C:Accession: S33788
 R:Deyne, J.H.; Kutuzova, G.D.; Green, V.A.; Ugarova, N.N.; Baldwin, T.O.
 Biochim. Biophys. Acta 1173, 121-132, 1993
 A:Title: Luciferase from the East European firefly Luciola mingrellica: cloning and nu
 A:Reference number: S33788; MUID:93277944
 A:Accession: S33788
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-548 <DEY>
 A:Cross-references: GB:S61961; NID:9409316; PIDN:AAB26932.1; PID:9409317
 C:Superfamily: 4-comumate--CoA ligase; acetate--CoA ligase homology
 C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase
 F:72-537/Domain: acetate--CoA ligase homology <ACL>

Query Match 82.7%; Score 2335.5; DB 2; Length 548;
 Best Local Similarity 81.7%; Pred. No. 1,4e-158;

Matches 446; Conservative 47; Mismatches 52; Indels 1; Gaps 1;

QY 4 MENMENDENIVYGPPEPPYPIEESGAGAOILRKYMDRYAKGALAFNALGVDTYAEYLEKSC 63
 |||||
 Db 3 MEKENYVYGPPLPPYPIEESGAGIOLKHYAKGALAFNALGVDTYAEYFDITC 62
 QY 64 CLGEALKNGLVVDGRIALCSENCEFFIPVLAGLFGVAPVNEIYTLRELHSLIGIS 123
 |||||
 Db 63 RLAMKMFQKDEBHALCSENCEFFIPVLAGLFGVAPVNEIYTLRELHSLIGIS 122
 QY 124 KPTIVFSSKKGLDKVITVQKTAIKTIYIIDSQKDYRGYQSDNFIKKNPQGGKSSF 183
 |||||
 Db 123 OPTIVFSSRRGLPKVLEVQKTIKTIYIIDSQKDYRGYQSDNFIKKNPQGGKSSF 182
 QY 184 KTEEV-NRKEQVALINSSSGSTGLPKGYOLTHEMLVTRSHADPIYGNQVSGTAVL 242
 |||||
 Db 183 VPLDVKKRQHVALLINSSSGSTGLPKGYOLTHEMLVTRSHADPIYGNQVSGTAVL 242
 QY 243 VPHHGFMTTGLYTCGFRIVMLTKPDEETFLKTLQDYKCSSVILVPTLAILNRS 302
 |||||
 Db 243 VPHHGFMTTGLYTCGFRIVMLTKPDEETFLKTLQDYKCSSVILVPTLAILNRS 302
 QY 303 LDKYLSNLVEIASGAPLSKEIGAVARFNPGLRGVGLTETTSAILITPEGDGK 362
 |||||
 Db 303 LDKYLSNLVEIASGAPLSKEIGAVARFNPGLRGVGLTETTSAILITPEGDGK 362
 QY 363 ASKGVVPLFKAKVIDLDTKTLGPNRRGECVCKGPMKGVNDPEARETIIDEEGWL 422
 |||||
 Db 363 ASKGVVPLFKAKVIDLDTKTLGPNRRGECVCKGPMKGVNDPEARETIIDEEGWL 422

QY	423	GDIGYDEKHEFFIYDRKLSLITKGYOVPAELESYLLQHPNIFDAGVAGVPDIAGEL	482
Db	423	GDIGYDEKHEFFIYDRKLSLITKGYOVPAELESYLLQHPNIFDAGVAGVPDIAGEL	482
QY	483	PGAVVLLTKGKSMTEKEVMDVVASOVSNAKRLRGVRFVDEVPKGLTQKIDGKATRETLK	542
Db	483	PGAVVVMKKGKTEKEVMDVVASOVSNVHKKRLRGVRFVDEVPKGLTQKIDAKVIRETLK	542
QY	543	KPVAKM	548
Db	543	KPVAKM	548

RESULT 4
S62787
Photinus-luciferin 4-monoxygenase (ATP-hydrolyzing) (EC 1.13.12.7) - Lampyris noctiluca
N:Alternate names: firefly-type luciferase
C:Species: Lampyris noctiluca
C>Date: 23-Aug-1996 #sequence_revision 01-Nov-1996 #text-change 26-May-2000
C:Accession: S62787; S57949
R:Salal-Newby, G.B.; Thomson, C.M.; Campbell, A.K.
B:Biochem. J. 313, 761-767, 1996
A:Title: Sequence and biochemical similarities between the luciferases of the glow-worm
A:Reference number: S62787; MUID:96190714
A:Accession: S62787
A:Molecule type: mRNA
A:Residues: 1-547 <SAL>
A:Cross-references: EMBL:X89479; NID:G899314; PIDD:CA61668.1; PID:G899315
C:Superfamily: 4-compartate--COA ligase; acetate--COA ligase homology
C:Keywords: ATP; luminescence; monoxygenase; oxidoreductase
P:71-535/Domian: acetate--COA ligase homology <ACLD>

Query Match	69.6%	Score 1965.5	DB 2	Length 547
Best Local Similarity	69.0%	Pred. No. 3.2e-132		
Matches 372	Conservative 76	Mismatches 90	Indels 1	Gaps 1
QY	4	MENDENITVGGPEPPYPIIEGSGAQLRKRYMDYAKL-GALAFNALTGVDYTAEYLKRS	62	
Db	1	MEDAKNIMHGPAFPYPLEDGTAGEQLHKMKRYAQPCTIAFTDAHAENVITSEYFEMA	60	
QY	63	CCGGEALKNKTVLVDGKHALCSENCCEFFTPVLAGLPIGVAAPTNNEYTTRELVHSGI	122	
Db	61	CRLAETMKRGLGLOHHIIVACSENSLOPFMPVGCALPIGVAASTNDINDNRELYNSLSI	120	
QY	123	SKPIIVFSSKKGLDVTYVTKYTAIKTIVLIDSKVDGRQYSDMNFJKKTPPOGFKSS	182	
Db	121	SGPIIVSCSKRALOKIIGVQKKLPIDQIVILSDREDIMQOSMTFLESILPAGFNFD	180	
QY	183	FKTIVENRKQVALLINMSGSTGLPKGVOLTHENLVTRFSHARDPIYGNQVSPGTAITLV	242	
Db	181	YIPDSFPERATAPALINMSGSTGLPKGVOLTHQVCRFSRCDPVPNGQIIPPTALITV	240	
QY	243	VPFHNGGMTTLLGYLLCGRIYMLTFDETFELKTLQDYKCSSVILPVLFAILNSEL	302	
Db	241	IPFHNGGMTTLLGYLLCGRIYMLTRFEELFLRSLODYVIOGALLVPTLFSFAKSTL	300	
QY	303	LDKKDLNLVEIASGGAPLSKEIGEAVARBNLPGVQGVGLTETSALITTPGDDKPG	362	
Db	301	VDKTDLSNLHEIASGGAPLKEVGEAVAKRFKLGLGIRQGGVLTETTSALITTPGDDKPG	360	
QY	363	ASGVVPLFRAKYIDLTPTKTLGPNRREYCVKSGPMLIKGVDMPEATREIIDEWGLHT	422	
Db	361	ACGVVVEFFSAKIYDLDPTGTLGVNQSGELCVKGPMLIKGVNNPEATSAIIDDGWLHS	420	
QY	423	GDIGYDEEKNHFFVYDKLSLITKGYQVPALEESVLLQHPNIFDAGVAVPPIAGEL	482	
Db	421	GDIAIYDKDHFIVYDKLSLITKIKGYQVPALEESILLQHPFIEDGAVAGIIPPDAGEL	480	
QY	483	PGAAVVLKKGKSMERKVMYVVAQVSNARKRGVGFVDEVPGLGKIDGKAIREL	541	
Db	481	PAAVVLEBGKTMDEYEMDVAGQVASKRLRGVGFVDEVPGLGKIDGKAIREL	539	

RESULT 5
 A26772
 Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) - common eastern
 N:Alternate names: firefly luciferase
 C:Species: Photinus pyralis (common eastern firefly)
 C:Date: 19-Nov-1988 #sequence_revision 20-Aug-1994 #text_change 26-May-2000
 C:Accession: A26772
 R:De Wet, J.R.; Wood, K.V.; Deluca, M.; Helinski, D.R.; Subramani, S.
 M:1. Cell. Biol. 7, 725-737, 1987
 A:Title: Firefly luciferase gene: structure and expression in mammalian cells.
 A:Reference number: A26772; MUID:87144243
 A:Accession: A26772
 A:Molecule type: DNA
 A:Residues: 1-550 <DEV>
 A:Cross-references: GB:MI5077; NID:G160793; PIDN:AA29795.1; PID:G160794
 A:Note: the authors translated the codon CAA for residue 134 as Glu
 C:Comment: This protein catalyzes the oxidation of luciferin in the presence of ATP,
 pecties from green to yellow.
 C:Genetics:
 A:Introns: 42/3; 112/1; 223/2; 330/2; 448/3; 504/1
 C:Superfamily: 4-commarate--CoA ligase; acetate--CoA ligase homology
 C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase
 F:71-535/Domain: acetate--CoA ligase homology <ACI>
 F:548-550/Region: peroxisome/glyoxysome location signal (S-[RKH]-L) motif

Query Match	68.9%	Score 1945.5	DB 1	Length 550
Best Local Similarity	68.2%	Pred. No. 8.6e-131		
Matches 369	Conservative 75	Mismatches 96	Indels 1	Gaps .1

[illegible]

RESULT 6

F:70-532/Domain: acetate--CoA ligase homology <ACL>

Query Match 48.2%; Score 1361; DB 2; Length 543;

Best Local Similarity 48.2%; Pred. No. 3, 6e-89;

Matches 262; Conservative 109; Mismatches 167; Indels 6; Gaps 3;

QY 4 MENDENIYVGPPEPPYIEEGSAGAO LRKYMDRYAKLGAIAFTNALTGVDYTAAYILEKSC 63
 DB 2 MKREKNVYGPPEPLHLEDTAGEMFLRALKHSHLPQ-ALVDVFDDESISYKFEFEAC 60
 QY 64 CLEBAKKNVGLVVDGRIALCSECEFFIPVLAGLPIGVAPNTNLTRELVSHLSIGIS 123
 DB 61 LLAQSLHNCYKMDVYSICAEENKREFPIITAMTIGMVAIVANSYIPDELCKMGIS 120
 QY 124 KPTIVSSKRGDLRVITVOKTVAIKTIVLDSKVDYRGYOSDNFKKNTPOGFGSSSF 183
 DB 121 KPTIVFCTKMLNKVLEVOGRTPFKRIIILDVENIHGCESLPNFTISRYSDGNI--ANF 178
 QY 184 KYEVNKRKEOVALLMNSGSGTGLPKGYQLTHEMLVTRFSHARDPIYGNVSPGTALITV 243
 DB 179 KPLHYDVEQVAAILCSSGTTGLPKGYQTHQNIQVRLHALDPEAGTOLIPGVTLVLY 238
 QY 244 PEHNGFMTTGLVTCGFRIVMLTKPDEETPLKTLQDYKSSVILVPTLFAILNSEL 303
 DB 239 PFHAFPGINLGVFMGLKVMIRRFQEAFLKAIQDYEVRSIVNPAIILFLSKSPV 298
 QY 304 DKYDLSNLVETASGAPLSKEIGEAARFNLPGVROGYGLTETSIIITPEGDDPGA 363
 DB 299 DKYDLSNLRELCCGAPLAKAEVAIAVKRLNLPICRGFGITESTSANIHSIDDEFSGS 358
 QY 364 SGKVPVLFKAKVIDLDTKTLGPRRRENVCKGPMKMGVNDPEAREIIDEQWHLTG 423
 DB 359 LGRYTPMAAKIADRETGKALGPQVGEELCTKGPWSKGYNNVAKAKELIDDDGWLHSG 418
 QY 424 DIGYDEKHEFFIYDRKLSLTKYKGYVPPAELESVLLQHPNIFDAGVAGVPPPIAGELP 483
 DB 419 DFGYDDEHHYVDRKELIKYKGOVAPAELEILLKNCIRDAVAVGIPDLAEGLP 478
 QY 484 GAVVYLKKGSMTEKEMVDYASOVSNARKRGVRFVDEVPKGLTKIDGKAIKRELTK 543
 DB 479 SAFVVIQPKKEITAEKVYDYLAERVSHTKYLRGVRFVDSIPRNVTKGKIRK--ELLKQ 535
 QY 544 PVAK 547
 DB 536 LLEK 539

RESULT 9

S29352
 Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) GR [validated] - Lum
 N:Alternate names: firefly-type luciferase
 C:Species: Pyrophorus plagiophthalmus
 C:Date: 22-Nov-1993 #sequence_revision 01-Sep-1995 #text_change 02-Jun-2000
 C:Accession: S29352
 R:Wood, K.V.; Lam, Y.A.; Seliger, H.H.; McElroy, W.D.
 Science 244, 700-702, 1989
 A:Title: Complementary DNA coding click beetle luciferases can elicit bioluminescence of
 A:Reference number: S29352; MUID:89242142
 A:Accession: S29352
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-543 <MOO>
 C:Superfamily: 4-coumarate--CoA ligase: acetate--CoA ligase homology
 C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase
 F:70-532/Domain: acetate--CoA ligase homology <ACL>

Query Match 48.0%; Score 1354; DB 2; Length 543;
 Best Local Similarity 48.4%; Pred. No. 1, 1e-88;
 Matches 265; Conservative 105; Mismatches 164; Indels 14; Gaps 4;

QY 4 MENDENIYVGPPEPPYIEEGSAGAO LRKYMDRYAKLGAIAFTNALTGVD---YTAAYEL 59

DB 2 MKREKNVYGPPEPLHLEDTAGEMFLRALKHSHL-----PQALVDVYGGEWISYKFEF 56

QY 60 EKSCGLGALKNVGLVVDGRIALCSECEFFIPVLAGLPIGVAPNTNLTRELVSHLS 119

DB 57 EYTCLLAQLHNCYKMDVYSICAEENKREFPIITAMTIGMVAIVANSYIPDELCKMGIS 116

QY 120 LGISKPTIVSSKRGDLRVITVOKTVAIKTIVLDSKVDYRGYOSDNFKKNTPOGFG 179

DB 117 MGISRPLVCTKMLNKVLEVOGRTPFKRIIILDVENIHGCESLPNFTISRYSDGNI- 175

QY 180 GSFYTVENKRKEOVALLMNSGSGTGLPKGYQLTHEMLVTRFSHARDPIYGNVSPGTAL 239

DB 176 -ANFPLHNDPEQVAAILCSSGTTGLPKGYQTHQNIQVRLHALDPEAGTOLIPGVTL 234

QY 240 LTVVPPHNGFMTTGLVTCGFRIVMLTKPDEETPLKTLQDYKSSVILVPTLFAILN 299

DB 235 LVYLPFFHAFGFSINLGVFMGLKVMIRRFQEAFLKAIQDYEVRSIVNPAIILFLSK 294

QY 300 SELDKYDLSNLVETASGAPLSKEIGEAARFNLPGVROGYGLTETSIIITPEGDD 359

DB 295 SPLVDKTYDLSRELCCGAPLAKAEVAIAVKRLNLPICRGFGITESTSANIHSIDDEF 354

QY 360 KPGASGVVPLFKAKVIDLDTKTLGPRRRENVCKGPMKMGVNDPEAREIIDEQWHL 419

DB 355 KSGSLGRVTPMAAKIADRETGKALGPQVGEELCTKGPWSKGYNNVAKAKELIDDDGW 414

QY 420 LHTGIDGYDEKHEFFIYDRKLSLTKYKGYVPPAELESVLLQHPNIFDAGVAGVPP 479

DB 415 LHSGDGYDEDEHFFYVDRKELIKYKGOVAPAELEILLKNCIRDAVAVGIPDLAE 474

QY 480 GELPGAVVYLKKGSMTEKEMVDYASOVSNARKRGVRFVDEVPKGLTKIDGKAIRE 539

DB 475 GELPSAFVVIQPKKEITAEKVYDYLAERVSHTKYLRGVRFVDSIPRNVTKGKIRK--E 531

QY 540 ITRKPVAK 547

DB 532 LTKQLLEK 539

RESULT 10

S01667
 4-coumarate--CoA ligase (EC 6.2.1.12) (clone pc4CL-1) - parsley
 N:Alternate names: 4-coumaroyl-CoA synthetase
 C:Species: Petroselinum crispum (parsley)
 C:Date: 30-Sep-1989 #sequence_revision 20-Aug-1994 #text_change 05-May-2000
 C:Accession: S01667
 R:Lozoya, E.; Hoffmann, H.; Douglas, C.; Schulz, W.; Scheel, D.; Hahlbrock, K.
 Eur. J. Biochem. 176, 661-667, 1988
 A:Title: Primary structures and catalytic properties of isoenzymes encoded by the two
 A:Reference number: S01667; MUID:89005119
 A:Accession: S01667
 A:Molecule type: mRNA
 A:Residues: 1-544 <LOZ>
 A:Cross-references: EMBL:X13324; NID:g20431; PIDD:CAA31696.1; PID:g20432
 C:Genetics:
 A:Gene: 4CL-1
 C:Superfamily: 4-coumarate--CoA ligase: acetate--CoA ligase homology
 C:Keywords: acid-thiol ligase; coenzyme A; flavonoid biosynthesis
 F:74-533/Domain: acetate--CoA ligase homology <ACL>

Query Match 27.2%; Score 768.5; DB 1; Length 544;
 Best Local Similarity 36.2%; Pred. No. 5, 9e-47;
 Matches 187; Conservative 88; Mismatches 189; Indels 53; Gaps 10;

QY 46 NALTGVYTYAAYIEKSCGLGALKNVGLVVDGRIALCSECEFFIPVLAGLPIGVTA 105
 DB 47 NGATGEFFYSQVELLSRKYASGLNKLIGIOGDTIMLLPNSPEYFPALGASRYAIST 106
 QY 106 PTNFIYTLRELVSHLSIGSKPTIVSSKRGDLRVITVOKTVAIKTIVL---DSKYDYG 162
 DB 107 MANPFTISAVIYNLQKNSQAKLIITGACVYDKV---KQVAAKKNQIICIDAPQDCJH 162

```

QY 163 YOSMDIFKKNPQGGKSSFTVEVNRKREQVALLINSSSTGLPKGVOLTHENLVTR-- 220
Db 163 FSKLMADESEMP-----VWINSDDVAL-PYSSGTTLPKGMVLTCHGLVTSVA 212
QY 221 -----FSHARDPIYGNVSPGTAITLVPPFHHGFMFTLLGLNCGFR----IYM 266
Db 213 QGVDDGNPMLYMHSD-----VMCIPLERH---IYSLNAVLCCLAGAVTILI 258
QY 267 LTKFDEETFLTKLÖDYKCSSVILVPTLFAILNSELIDKYDSLNLVEIASGAPLSKEIG 336
Db 259 MÖKFIVPTELTELQKKVITIGFVPIVLAIAKSPYVDKLDSSVRYMSGAPLKEKE 318
QY 327 EAVARFNLPGVRGOGGLTETTSALII-----TPBEDDKPGASGKVVPLFKAVVIDLDTK 361
Db 319 DAVRAKFPMAKGGÖGGMTEAGPVLAMCLAFKPEYEITSGAGTIVRAEMKIVDPETN 378
QY 382 KTLGNRRGCEVCVKGMPLMKRGVNDPEATREIIDEKGMLHTGDIGYDEKHEFIYDRK 441
Db 379 ASLPNRÖGSEIIRBGÖIMKGYLNDPESTRITIIDEGWMLHTGDIGITDDDELFIYDRK 438
QY 442 SLITKKGIOVPAELLESVLLQHPNIFDAGVAGVPPDIAGELPGAAVVLLKKGSMTEKEYM 501
Db 439 ELIKKKGEOVAAELEALLTHPTISDAVVMIDEBKAGEVPAVAVVRNNGFTTDEELK 498
QY 502 DYVASOVMAKRLRGVRFVDEVPGKLGKIDOKAIR 538
Db 499 ÖPVSÖÖVVFYKRI-FRVFEVDALPKSSGKILIRKDIR 534

```

```

RESULT 11
S15695
4-comumarate--CoA ligase (EC 6.2.1.12) (clone Pc4CL-2) - parsley
C:Species: Petroselinum crispum (parsley)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-May-2000
C:Accession: S15695
R:Lozoya, E.; Hoffmann, H.; Douglas, C.; Schulz, W.; Scheel, D.; Hahlbrock, K.
Eur. J. Biochem. 176, 661-667, 1988
A>Title: Primary structures and catalytic properties of isoenzymes encoded by the two 4
A:Reference number: S01667; MUID:89005119
A:Accession: S15695
A:Molecule type: mRNA
A:Residues: 1-544 <LO2>
A:Cross-references: EMBL:X13325; NID:g20435; PIDN:CAA31697.1; PID:g20436
C:Genetics:
A:Gene: 4CL-2
C:Superfamily: 4-comumarate--CoA ligase; acetate--CoA ligase homology
C:Keywords: acid-thiol ligase; coenzyme A
F:74-533/Domain: acetate--CoA ligase homology <ACL>

```

[illegible]

Db	259	MOKFDVPLELLEOKKVVIGFVPPIVLAIKSPVYDKLIDLSVRTWMSGANPLQKELE	318
QY	327	EVAREFNLPGVRCQYGLTETTSALII-----TPBGDCRPGASGVVPLFKAAYIDLDTK	381
Db	319	DVRAFEFPAKLGQGYMTAGPVLAMCLAFAKEPEIKSGACGYVVRNAEKKIYDPETN	378
QY	382	KTLGPNRRGEVVCVKPMLKGYVNDPEATRELIIDEGWMLHTGIDIGYDYEKKHFTVDRK	441
Db	379	ASLPNNQREICIRDDQIMKGLNDEPSTRITIIDEGMLHTGIDIGFIDDDLELFTVDRK	438
QY	442	SLIKTKGIVPPAELESLVLLQHPNIFDAGVAGVPDPIAGELPGAAVVLKKGSKMTEEYM	501
Db	439	ELIKYKGEVPAAELEALILTHPTISDAAVVPMIDEKAGEVPAVVRNNGFTTTEEEK	498
QY	502	DYVASOVNAKRLRGVRFVDEVPKGLGKIDGKAIR	538
Db	499	QFVSKQVVFYKRI-FRVEFVDALPKSPSGKILIRKRO	534

```

RESULT 12
T07909
4- coumarate--CoA ligase (EC 6.2.1.12) 1 [validated] - western balsam poplar x cottonw
C:Species: Populus trichocarpa x Populus deltoides [western balsam poplar x cottonwoc
C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 26-May-2000
C:Accession: T07909
R:Allina, S.M.; Pri-Idash, A.; Theilmann, D.A.; Ellis, B.E.; Douglas, C.T.
Plant Physiol. 116, 743-754, 1998
A>Title: 4-Coumarate:Coenzyme A ligase in hybrid poplar. Properties of native enzymes
A:Reference number: Z16208; MUID:98150279
A:Accession: T07909
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-557 <ALL>
A:Cross-references: EMBL:AF008184; NID:g2911798; PIDN:MAC39366.1; PID:g2911799
A:Genetics:
A:Gene: 4CUL1
A:Function:
A:Description: EC 6.2.1.12; provides activated thioester substrates for phenylpropano
C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology
C:Keywords: acid-thioester ligase; coenzyme A; flavonoid biosynthesis
F:74-534/Domain: acetate--CoA ligase homology <ACLD>

```

[illegible]

Db 381 SUPRNOSGECICISQIMKCYLNDPEATEKTVNDGMLHTGDIYDGDLEFIYDRLKE 440

QY 443 LIKYGQVPALESLVLLQHPIEDAGVAGVDPDPIAGELPGAVVVLKKGSKTEKEVMD 502

Db 441 LIKYGQVPALESLVLLQHPIEDAGVAGVAGVDPDPIAGELPGAVVVLKKGSKTEKEVMD 500

QY 503 YVASOVNAKRLRGVRFVDEVPKGLTGKIDGKAIR 538

Db 501 YISKOVYFYKRI-SRVFETFAIRKPSGKILRKDLR 535

RESULT 13

T07908

4-coumarate--CoA ligase (EC 6.2.1.12) 2 - western balsam poplar x cottonwood

C:Species: Populus trichocarpa x Populus deltoides (western balsam poplar x cottonwood)

C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 11-May-2000

C:Accession: T07908

R:Allina, S.M.; Pri-Hadash, A.; Thellmann, D.A.; Ellis, B.E.; Douglas, C.J.

A:Title: 4-Coumarate:Coenzyme A ligase in hybrid poplar. Properties of native enzymes, plant physiol. 116, 743-754, 1998

A:Reference number: Z16208; M01D:98130279

A:Accession: T07908

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-548 <ALL>

A:Cross-references: EMBL:AF008183; NID:92911796; PIDN:ABC39365.1; PID:g2911797

C:Gene: 4CL2

C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology

C:Keywords: acid-thiol ligase; coenzyme A; flavonoid biosynthesis

F:74-533/Domain: acetate--CoA ligase homology <AC1>

Query Match 26.8%; Score 755.5; DB 2; Length 548;

Best Local Similarity 36.2%; Pred. No. 5e-46;

Matches 187; Conservative 85; Mismatches 193; Indels 51; Gaps 12;

QY 46 NATGVDTYAEYLEKSCCIGCALKNYGLVVDGRALCSECEFFPIVLAGLFGVGA 105

Db 47 NGPFGDITTYADVDELTSKRVASGLKIGLOGDVILLIIONSPFVAFGLASTIGAISS 106

QY 106 PTNETYTLRELVLHSLGSKPTIVSSKRGDLKVTYVOKTVTAIKTVIILDSKVDYRGYOS 165

Db 107 TANPFYTSAEIAKQATASKALITTHAAVAKYQOFAQENDHVK-IMTISLT----- 158

QY 166 MDNFKKTPGCGFKSGSKTYEVNRKQVALINMSGSTGPKGVOLTHEMLYR----- 220

Db 159 -ENCLHFEELTSSDENEIPYKRI-KPDDIMALPYSSGTTGLPKGVMLTHKGLVHSAQOV 216

QY 221 -----FSHARDPIYGNQVSPGTAILTVVPFHGFGMFTTLGYTCGFR-----IVMLTK 269

Db 217 DGENPMLYFHERD-----VILCVLPFLH---IYSLNSVFLGLGRAGSAILVMOK 262

QY 270 FDEETFLKTLQDYKCSYILVPTLFAILNRSLELDKYDLSNLVELASGAPLSKEIGBAV 329

Db 263 FDTYSLMDLVQKYKVTIAPLPICLAIAKSPVVDQDLSIRTVLSGLAGLELBDTV 322

QY 330 ARFPLPGVR--QGYGLTETTSAILI-----TPBGDDKPGASGVVPLFKAKVIDLDTKK 382

Db 323 --RAKLPNAKLGQGYGMEAGVPVIAMLCLAFKPEPEIKSGAGVVRNAEMKIYDPTGE 380

QY 383 TLGPRREGVCGKPMKKGVDNPEATREIIDEGLMLHTGDIYDEKHEFFIVDRLSKLS 442

Db 381 SOPRKTEICIRGQIKMGKYLNDPEATEERTIDKGMWMLHTDIDIGIDDE-LFTVDRLKE 439

QY 443 LIKYGQVPALESLVLLQHPIEDAGVAGVDPDPIAGELPGAVVVLKKGSKTEKEVMD 502

Db 440 LIKYGQVPALESLVLLQHPIEDAGVAGVAGVDPDPIAGELPGAVVVLKKGSKTEKEVMD 500

QY 503 YVASOVNAKRLRGVRFVDEVPKGLTGKIDGKAIR 538

Db 500 YISKOVYFYKRI-SRVFETFAIRKPSGKILRKDLR 534

RESULT 14

H85064

4-coumarate--CoA ligase-like protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001

C:Accession: H85064

R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp

A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A:Reference number: A85001; M01D:20083488

A:Accession: H85064

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-544 <STO>

A:Cross-references: GB:NC_001268; NID:97267275; PIDN:CA81058.1; GSPDB:GN00140

C:Gene: AT4G05160

C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology

Query Match 26.7%; Score 755; DB 2; Length 544;

Best Local Similarity 34.0%; Pred. No. 5.4e-46;

Matches 178; Conservative 111; Mismatches 195; Indels 40; Gaps 12;

QY 42 IAFNNALTGVDYTAAYLEKSCCIGCALKNYGLVVDGRALCSECEFFPIVLAGLFG 101

Db 44 LAIADSDTDLSLTSOLKSAVARLAHGHRLGIRKNDVLIAPNSYCPFLCFLAVTAIG 103

QY 102 VGVAPTNETYTLRELVLHSLGSKPTIV-----FSSKRGDLKVTYVOKTVTAIKTVIIL 155

Db 104 GVFTFANPLYTVNEVSKQIKDSNPKIILISVNOLEFKIRGFDLPVVLGSKDVEIIPGSN 163

QY 156 SKVDYGYOSMNFIRKTPGCGFKSGSKTYEVNRKQVALINMSGSTGPKGVOLTHEMLYR 215

Db 164 SKI-----LSFDVNMELSEPV-----SEYFPVEIKOSDTALLY-SSGTTGSKGVELTHG 213

QY 216 N-----LVTRFSHARDPIYGNQVSPGTAILTVVPFHGFGM-FTTLGYTCGFRIVMLT 268

Db 214 NFTAASLMTMDQDLGGEVHG-----VFLCFLPFHFVHGLAVITVSLQGNMLVMA 266

QY 269 KEDETFPLKTLQDYKCSYILVPTLFAILNRSLELDKYDLSNLVELASGAPLSKEIGBA 328

Db 267 RFELEVLKNIKIEKFRVTHLVVPPVPLALSQSIYKFPDLSLKYIGSAAFLGKDLME 326

QY 329 VARRFPLPGV--ROGYGLTETTSAILIITPE--GDDKPGASGVVPLFKAKVIDLDTKK 384

Db 327 CGR--NIPNVLLMOGYGMEAGVPVIAMLCLAFKPEPEIKSGAGVVRNAEMKIYDPTGE 380

QY 385 GPNRREGVCGKPMKKGVDNPEATREIIDEGLMLHTGDIYDEKHEFFIVDRLSKLS 444

Db 385 PPNQGGIWRGNNMKGYNLPNPAKRETIIDKSSVHTGDLGFNDGMLVYVDRKELI 444

QY 445 KYRGQVPALESLVLLQHPIEDAGVAGVDPDPIAGELPGAVVVLKKGSKTEKEVMD 504

Db 445 KYRGQVPALESLVLLQHPIEDAGVAGVAGVDPDPIAGELPGAVVVLKKGSKTEKEVMD 500

QY 505 ASOVNAKRLRGVRFVDEVPKGLTGKIDGKAIR 538

Db 505 AKOVAPYKRLR-RVSPFLVPKSAGRI--LRRELVOVRSKM 544

RESULT 15

T03789

4-coumarate--CoA ligase (EC 6.2.1.12) 4CL2 - common tobacco

N:Alternate names: 4-coumaroyl-CoA synthetase

C:Species: Nicotiana tabacum (common tobacco)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000

C:Accession: T03789

R:Lee, D.; Douglas, C.J.

A:Title: Two divergent members of a tobacco 4-coumarate:coenzyme A ligase (4CL) gene

A:Reference number: Z15086; MUID:96416441
A:Accession: T03789
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-542 <LEF>
A:Cross-references: EMBL:U50846; NID:g1663723; PIDN:AAB18638.1; PID:g1663724
C:Gene: 4CL2
C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology
C:Keywords: acid-thiol ligase; coenzyme A; flavonoid biosynthesis
F:73-532/Domain: acetate--CoA ligase homology <ACL>

Query Match 26.5%; Score 748.5; DB 2; Length 542;
Best Local Similarity 35.1%; Pred. No. 1.6e-45;
Matches 190; Conservative 89; Mismatches 200; Indels 63; Gaps 12;

QY 32 YMDRYAKGIAIAFTN-----ALTGVD---YTYAEYLEKSCCLGEALKNYGLVVDGR 79
DB 20 YIPNHLPDLSHYCFENISESSRPCLINGANKQIYTADELNSRKVAAGLHKGQIQPKDT 79
QY 80 IALSCENCEFFPIPVLAGLFIQVAPFTEIYTLRELHSLGISKPTIYFSSKGLDKVI 139
DB 80 IMILPNSPEFVAFIAGSYLAGIISMANPLFPAEVVQAKASSAKIIVTQACHVANKV- 138
QY 140 TVQKTYTAKTIYIILDSKDYRGYQSMDFIK---KNTPOGPKSSSEFT-----V 186
DB 139 -----KDY---AFENDVKIICIDSAPEGCLHFSVLTQANEHDIPEV 176
QY 187 EYNRKEQVALIMNSSGSGTGLPKGVOLTHENLVTRFSHARDPIYGN-QVSPGTAILTVVF 245
DB 177 ELQPPDDYVAL-PYSSGTITGLPKGVMLTHGLVTSVAQVQDGENPNLYIHSEVWMLCVLPL 235
QY 246 HHGFGMETTLGYLTGFR---IVMLTKFDETEFLTKLDYKCSSYILVPTLEPAILNRSE 301
DB 236 FH---IYSLNSVLTGCLRGAAILIMQKFDIVSELELIQRYKVTIGPFVPIVLAIAKSP 292
QY 302 LLDKYDLSMLVEIASGAPLSKEIGAVARRENLPGVROGYGLTETTSAILI-----TPE 356
DB 293 MVDYDLSVSVKTVMSGAPLGELEDTVRAKFPNNAKLGQGYGTEAGPYLAMCLAFAPKEP 352
QY 357 GDDKPGASGVVPLFAKAVIDLTKTKLGNRRGEVGVKPMIMKGYVNDPEATREIIDE 416
DB 353 FEIKSGAGCTVVRNEMKIVDPKTGNSLPRNGSEICIRGDIIMKGYLNDPEATARTIDK 412
QY 417 EGMULTGDIQYDEKHFIVRLKSLIKYKGYQVPAELESVLLQHPNIFDAGVAVPD 476
DB 413 EGMULTGDIQYIDDDDELIVRLKELIKYKGYVAPAELEALLNHPNISDAAVVPMKD 472
QY 477 PIAGEIPGAVVVLKKGKSMTEKEVMDYVASOVSNAKRLGVRFVDEVPKGLTKIDGKA 536
DB 473 EQAGEVPVAFVVRNSGSTITTEDEVAKDFISKQVIFYKRIK-RVFFVDATPKSPSGKILRKD 531
QY 537 IR 538
DB 532 LR 533

Search completed: November 29, 2001, 01:57:51
Job time: 361 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2001, 01:57:51 ; Search time 36.66 Seconds
(without alignments)
1138.670 Million cell updates/sec

Title: US-09-581-241-6

Sequence: 1 MNNNDENIVYGEPEPPYPL.....TGKIDKAIKREILKKPYAKM 548

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR:68:*
2: PIR:1:*
3: PIR:3:*
4: PIR:4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2814	99.7	548	1 S23437	Photinus-luciferin
2	2684	95.1	548	1 JS0181	Photinus-luciferin
3	2335.5	82.7	548	1 S33788	Photinus-luciferin
4	1967.5	69.7	547	2 S62787	Photinus-luciferin
5	1945.5	68.9	550	1 A26772	Photinus-luciferin
6	1367	48.4	543	2 S29354	Photinus-luciferin
7	1367	48.4	543	2 S29353	Photinus-luciferin
8	1363	48.3	543	2 S29355	Photinus-luciferin
9	1356	48.0	543	2 S29352	Photinus-luciferin
10	766.5	27.2	544	1 S01667	4-comumate--CoA 1
11	758.5	26.9	544	2 S15695	4-comumate--CoA 1
12	756	26.8	557	2 T07909	4-comumate--CoA 1
13	754	26.7	544	2 H85064	4-comumate--CoA 1
14	753.5	26.7	548	2 T07908	4-comumate--CoA 1
15	746.5	26.4	542	2 T03789	4-comumate--CoA 1
16	738.5	26.2	542	2 T02074	4-comumate--CoA 1
17	732.5	25.9	535	2 T08074	4-comumate--CoA 1
18	732.5	25.9	545	1 A39827	4-comumate--CoA 1
19	731	25.9	561	2 D96874	hypothetical prote
20	730	25.9	563	1 J00311	4-comumate--CoA 1
21	725.5	25.7	545	2 B39827	4-comumate--CoA 1
22	720	25.5	544	2 T20741	hypothetical prote
23	719	25.5	549	2 D86338	protein F5M15.18
24	718.5	25.5	561	2 S57784	4-comumate--CoA 1
25	713	25.3	570	2 T08075	4-comumate--CoA 1
26	691.5	24.5	569	2 T03390	4-comumate--CoA 1
27	680	24.1	537	2 T09755	4-comumate--CoA 1
28	674.5	23.9	542	2 B86654	hypothetical prote
29	672	23.8	537	2 T09710	4-comumate--CoA 1

30	659.5	23.4	423	2 P00772	4-comumate--CoA 1
31	635	22.5	552	2 E69438	probable fatty-aci
32	632.5	22.4	566	2 F85214	4-comumate--CoA 1
33	632.5	22.4	566	2 T05038	4-comumate--CoA 1
34	600	21.3	566	2 H84037	long-chain fatty-a
35	597.5	21.2	569	2 C69471	probable fatty-aci
36	587.5	21.1	564	2 G84037	long-chain fatty-a
37	587.5	20.8	293	2 S31705	4-comumate--CoA 1
38	581.5	20.6	593	2 E69378	probable acid--CoA
39	577	20.4	548	2 D69187	probable acid--CoA
40	574.5	20.4	560	2 D69649	probable long-cha
41	573	20.3	540	2 T18841	hypothetical prote
42	568	20.1	561	2 S41589	long-chain-fatty-a
43	568	20.1	561	2 F85791	hypothetical prote
44	565	20.0	546	2 C83791	acid-CoA ligase BH
45	564.5	20.0	584	2 C75364	probable long-cha

ALIGNMENTS

RESULT 1

S23437
Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) - Japanese firefly
N:Alternate names: firefly luciferase
C:Species: Luciola lateralis (Japanese firefly)
C:Date: 22-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 02-Jun-2000
C:Accession: S23437; S57417
R:Tatsumi, H.; Kajiyama, N.; Nakano, E.
Biochim. Biophys. Acta 1131, 161-165, 1992
A:Title: Molecular cloning and expression in *Escherichia coli* of a cDNA clone encodin
A:Reference number: S23437; MUID:92305054
A:Accession: S23437
A:Molecule type: mRNA
A:Residues: 1-548 <TRAIL>
A:Cross-references: EMBL:X66919; NID:99526; PIDN:CAA47358.1; PID:99527
R:Cho, K.; Choi, Y.; Boe, K.
Submitted to the EMBL Data Library, June 1995
A:Description: Molecular cloning of gene for luciferase in *Luciola lateralis*.
A:Reference number: S57417
A:Accession: S57417
A:Molecule type: DNA
A:Residues: 1-4,'D',6-14,'K',16-145,'N',147-175,'P',177-505,'G',507-548 <CHOW>
A:Cross-references: EMBL:Z49991; NID:9871400; PIDN:CAA90072.1; PID:9871401
C:Genetics:
A:Introns: 44/3; 114/1; 225/2; 332/2; 450/3; 506/1
C:Function:
A:Description: catalyzes the irreversible reaction of luciferin with oxygen and ATP to
C:Superfamily: 4-comumate--CoA ligase; acetate--CoA ligase homology
C:Keywords: ATP; luminescence; magnesium; monooxygenase; oxidoreductase; peroxisome
F:73-537/Domain: acetate--CoA ligase homology <ACLT>

Query Match Score 2814; DB 1: Length 548;

Best Local Similarity 99.6%; Pred. No. 2,7e-192;
Matches 546; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY	1	MNNNDENIVYGEPEPPYPIEESAGAKYMDRAKALGALFTNALGVDTVAEYLE 60
DB	1	MNNNDENIVYGEPEPPYPIEESAGAKYMDRAKALGALFTNALGVDTVAEYLE 60
QY	61	KSCGLGALKNYGLVNDGRIALCSNCEEFFIVLAGLFGVAPPTNEITLRELVSHL 120
DB	61	KSCGLGALKNYGLVNDGRIALCSNCEEFFIVLAGLFGVAPPTNEITLRELVSHL 120
QY	121	GISKPTVFSKKGKGLKVITVOKTVAIKTIVLDSVDVGRGVSDMNFKKTPPGFG 180
DB	121	GISKPTVFSKKGKGLKVITVOKTVAIKTIVLDSVDVGRGVSDMNFKKTPPGFG 180
QY	181	SSFTVFNRRKEQVALIMNSGSGTGLPKGVOLTHENIVTRFSSHARDPIYGNVSPGTAIL 240
DB	181	SSFTVFNRRKEQVALIMNSGSGTGLPKGVOLTHENIVTRFSSHARDPIYGNVSPGTAIL 240

Oy	241	TVPFHHGGMFTLLYLTCGFRIVMLTKFEDEETELKTIDYKCCSVILVPTLALLNRS	300
Dd	241	TVPFHHGGMFTLLYLTCGFRIVMLTKFEDEETELKTIDYKCCSVILVPTLALLNRS	300
Oy	301	ELLDKYDLSNLVEIASGAPPLSKKEIGEAARRENLPGRVGGLTFETTSAIIITPEGDCK	360
Dd	301	ELLDKYDLSNLVEIASGAPPLSKKEIGEAARRENLPGRVGGLTFETTSAIIITPEGDCK	360
Oy	361	PASGCVVPLFLAKVIDLDTKTTLGPNRGEVCVKGPMLMKGYVDNPENATRETIDEEGLW	420
Dd	361	PASGCVVPLFLAKVIDLDTKTTLGPNRGEVCVKGPMLMKGYVDNPENATRETIDEEGLW	420
Oy	421	HFGDIGYDEEHFFIYVDRLKSILIKKYQVPPAELESYLLOHPNIFDAGVGVDPPIAG	480
Dd	421	HFGDIGYDEEHFFIYVDRLKSILIKKYQVPPAELESYLLOHPNIFDAGVGVDPPIAG	480
Oy	481	ELPGAVVLKKKSMTEKEVM DYVASQVSNAKRRLRGVAFVDEVPKGLGKIDGAIRRI	540
Dd	481	ELPGAVVLLEKSKSMTKEVM DYVASQVSNAKRRLRGVAFVDEVPKGLGKIDGAIRRI	540
Oy	541	LKKPVAKM 548	
Dd	541	LKKPVAKM 548	

RESULT 2

```

JS0181
Protein: luciferin 4-monoxygenase (Atp-hydrolyzing) (EC 1.13.12.7) - GenJl firefly
N:Alternate names: firefly luciferase
C:Species: Luciola cruciata (GenJl firefly)
C>Date: 07-Jun-1990 #sequence_revision 20-Aug-1994 #text_change 26-May-2000
R:Accession: JS0181
R:Masuda, T.; Tatsumi, H.; Nakano, E.
Gene 77, 265-270, 1989
A>Title: Cloning and sequence analysis of cDNA for luciferase of a Japanese firefly, Luc
A:Reference number: JS0181; MUID: 89326143
A:Accession: JS0181
A:Molecule type: mRNA
A:Residues: 1-548 <MAS>
A:Cross-references: GB:M6194; NID:g159050; PIND:AAA29135.1; PID:g159051
C:Comment: This protein catalyzes the oxidation of Luciferin in the presence of ATP, oxy
pecies from green to yellow.
C:Superfamily: 4-comarate--CoA ligase; acetate--CoA ligase homology
C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase
C:73-53//Domain: acetate--CoA ligase homology <ACID>
```

Db	301	ELLINKDLSNLVEIISGAPLRSKEWGEAAVRPNLPVQSGIGLETTSALITTEPEGDDK	360
QY	361	PGASGVVDFKAKYLDLDTKTKTLCGPNRGEVCVAGPMLMKGYVNPPEATREITDEEGL	420
Db	361	PGASGVVDFKAKYLDLDTKTSLSGPNRGEVCVAGPMLMKGYVNPPEATREITDEEGL	420
QY	421	HGIGDIGYDDEEHFFIVDLAKSLIKYGYQVPAELESVLLQHPNIFDAGVAGVPDPIAG	480
Db	421	HGIGDIGYDDEEHFFIVDLAKSLIKYGYQVPAELESVLLQHPNIFDAGVAGVPDPIAG	480
QY	481	ELPGAVVLLKKSSTTEREVMQVYASQVSNMRLKGGVRFDEVEYKGLTGKIDGAIETI	540
Db	481	ELPGAVVLLKKSSTTEREVMQVYASQVSNMRLKGGVRFDEVEYKGLTGKIDGAIETI	540
QY	541	LKKPVAKM 548	
Db	541	LKKPVAKM 548	

RESULT 3
S33788
Phospho-Luciferin 4-monooxygenase (AMP-hydrolyzing) (EC 1.13.12.7) [validated] - source
N:Alternate names: firefly luciferase
C:Species: *Luciola mingrelica* (Southern Russian firefly)
C:Date: 02-Dec-1993 #sequence_revision 03-Nov-1995 #text_change 02-Jun-2000
C:Accession: S33788
R:Devine, J.H.; Koltunova, G.D.; Green, V.A.; Ugarkova, N.N.; Baldwin, T.O.
Biochem. Biophys. Acta 1173, 121-132, 1993
A:Title: Luciferase from the East European firefly *Luciola mingrelica*: cloning and nu
A:Reference number: S33788; MUID:93277944
A:Accession: S33788
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-548 <DEV>
A:Cross-references: GB:S61961; NID:g409316; PIDN:AAB26932.1; PID:g4009317
C:Superfamily: 4-conmarate-COA ligase, acetate-COA ligase homology
C:Keywords: AMP, luminescence, monooxygenase, oxidoreductase
F:72-537/Domain: acetate-COA ligase homology <ACLD>

Query Match	95.1%	Score 2684	DB 1	Length 548
Best Local Similarity	93.4%	Pred. No. 4.7e-183		
Matches 512	Conservative 26	Mismatches 10	Indels 0	Gaps 0
QY	1	MENNENDENIYGEPEPTPIEGSAGAO LRKRYMDRYAKGAIATFNALTGVDYTAEYLE	60	
Db	1	MENNENDENIYVGPKEPPPIEGSAGTOLRKRYMERYAKGAIATFNNAVGTGVDTAEYLE	60	
QY	61	KSCCIGALKNKYGLVVDGRILALCSENCEEPFIPYIAGLFGTGVAPNNEYITTLAREVHSL	120	
Db	61	KSCCIGALKNKYGLVVDGRILALCSENCEEPFIPYIAGLFGTGVAPNNEYITTLAREVHSL	120	
QY	121	GISKPTIVSESKRKLDKVIYQKTVTLAKTIVILDSVDYIRGVOSMDNFIKKNTPOGFKG	180	
Db	121	GISKPTIVSESKRKLDKVIYQKTVTLAKTIVILDSVDYIRGVOCIDPTFKRNTPPGFOA	180	
QY	181	SSKPTVEVNRKEOVALIMNSGSGTGLPKGVOLTHENIVTFRSHARDPIYGNVSPGAVIL	240	
Db	181	SSKPTVEVDRKEOVALIMNSGSGTGLPKGVOLTHENIVTFRSHARDPIYGNVSPGAVIL	240	
QY	241	TVVPFHNGPMEFTTIGVLTGCGFRIVMLTKRDEDEFLKTLDDYKCSSVILVPTLPAIINRS	300	
Db	241	TVVPFHNGPMEFTTIGVLTGCGFRIVMLTKRDEDEFLKTLDDYKCSSVILVPTLPAIINRS	300	
QY	301	ELLDKDYLSNIVEIASGAPLSEKIEGAVARRNPLPVRGQYGTETTSATIIITPECDK	360	

Query Match	82.7%	Score 2335.5	DB 2	Length 548
Best Local Similarity	81.7%	Pred. No. 2,36	158	
Matches 446	Conservative 47	Mismatches 52	Indels 1	Gaps 1
QY	4	MENDENITVGEPEPPYPIEESGAGQLRKRYMDRYKALIAFTNALTGVDITYAEYLKSC	63	
Db	3	MEKENEVVYGPFLPEYPIEESGAGIQLHKYMHQYAKKALIAFNSALITGVDSIYQEPYIC	62	
QY	64	CLGALKNNYGLVYDGRILALCSNCEEFFIYVLAQLFTGCVAPLPNEIYTLRELVHSLGIS	123	
Db	63	RLADAMNFGKKRPEHIALALCSNCEEFFIYVLAQLFTGCVAPLPNEIYTLRELHNSLGIA	122	
QY	124	KPIIVSSSKKGLKLVITVQKTVTAIKTVILDSKVDYRGQSMDFPKKWTPOGFKSSP	183	
Db	123	QPIIVFSRRKGLPKVLEIVQKTVTCIKKIIVILDSKVNNGCHDCMETFLKKHVELGFPQSSP	182	
QY	184	KTYVY-NRKQVALIMNSSGSTGLPKGVOLTHENITYRSHADPIYGNOSGTALITV	242	
Db	183	VPIIDKKRKHQVALIMNSSGSTGLPKGVRLTHGCAVTRFSHADPIYGNOSGTALITV	242	
QY	243	VPHHNGMFTTLAGYLTCGFRIVMLTKFDETEFLKTLIDVCKSSVILVPLFLILRNSL	302	
Db	243	VPHHNGMFTTLAGYFACGYRVMLTKFDEDELRLTLIDYKCTSVILVPLFLILKNSL	302	
QY	303	LKRYDLSNLVEIASGAPLSKEIGEAVARRFNLPVROGYGLTETTSAILIIPEDGDKRG	362	
Db	303	IDFDSLNLTEIASGAPLAKEVEAVARRFNLPVROGYGLTETTSAILIIPEDGDKRG	362	
QY	363	ASQKVPFLFAKVIDTDTKTLTGPNRRGEVCGPMLKGYVNPATREIIDEBSWLHT	422	
Db	363	ASQKVPFLFAKVIDTDTKTLTGPNRRGELCYVGSFLMIGYSNNPATREIIDEBSWLHT	422	

QY 423 GDIGYDEEHFFIVDLRLSLIKYGYVPPALESVLLQHPNIDAGVAGVDPPIAGEL 482
 |||||
 Db 423 GDIGYDEEHFFIVDLRLSLIKYGYVPPALESVLLQHPNIDAGVAGVDPPIAGEL 482
 QY 483 PGAVVVLKKGKSMTEKEMDVYASQVSNARKLRGVRFEVDEVRKGLTGKIDGKAIREILK 542
 |||||
 Db 483 PGAVVVLKKGKSMTEKEMDVYASQVSNARKLRGVRFEVDEVRKGLTGKIDGKAIREILK 542
 QY 543 KPVAKM 548
 |||||
 Db 543 KPVAKM 548

RESULT 4
 S62787
 Photinus-luciferin 4-monoxygenase (ATP-hydrolyzing) (EC 1.13.12.7) - Lampyris noctiluca
 N:Alternate names: firefly-type luciferase
 C:Species: lampyris noctiluca
 C>Date: 23-Aug-1996 #sequence_revision 01-Nov-1996 #text_change 26-May-2000
 C:Accession: S62787; S57949
 R:Salal-Newby, G.B.; Thomson, C.M.; Campbell, A.K.
 Biochem. J. 313, 761-767, 1996
 A:Title: Sequence and biochemical similarities between the luciferases of the glow-worm
 A:Accession number: S62787; MUID:96190714
 A:Reference: S62787
 A:Molecule type: mRNA
 A:Residues: 1-547 <SAL>
 A:Cross-references: EMBL:X89479; NID:g989314; PIDN:CAA61668.1; PID:g989315
 C:Superfamily: 4-comumate--CoA ligase; acetate--CoA ligase homology
 C:Keywords: ATP; luminescence; monoxygenase; oxidoreductase
 F:71-535/Domain: acetate--CoA ligase homology <ACL>

Query Match 69.7%; Score 1967.5; DB 2; Length 547;
 Best Local Similarity 69.0%; Pred. No. 4.1e-132;
 Matches 372; Conservative 76; Mismatches 90; Indels 1; Gaps 1;

QY 4 MENDENIVYGEPEPYPIEESAGAOQLRKRYMDYAKL-GATAFTNALTGVDYTAELYEKS 62
 |||||
 Db 1 MEDAKNIHKGAPFYPLEDEGTAGEOLHKAMKRYAGVGTIAFTNAHEVNTIYEYEMA 60
 QY 63 CCLGEALKNYGLVVDGRIALCSECEFFIPVLAGLFTGVGAPATNEITYTLRELVHSLGI 122
 |||||
 Db 61 CRLEMTKRYGLGLOHNIHIVCSENSLOPFMPVCGALFVGAVASTNDIYNRELINSLSI 120
 QY 123 SKPTIVSSKGLDKVITVQVTAIKTIYVLDKVDYRGQSDNFIKKNTPGFGSS 182
 |||||
 Db 121 SOPTIVSSKRALOKILGVOKRPLIKRIYVLDSDREDYMGQSMVSELSHLPAGFNEYD 180
 QY 183 FKTVEVNRKDOVALIMNMSGSTGLPKGVOLTHENIVTRFSHARDPIYGNQVSPGAILTV 242
 |||||
 Db 181 YIPDSFDEETATALIMNMSGSTGLPKGVOLTHENIVTRFSHARDPIYGNQVSPGAILTV 240
 QY 243 VPFHNGFMFTTGLYTCGRFVLMFKFDEFTFLKTDYKCSVILVPTLPAIINSEL 302
 |||||
 Db 241 IPEFHGFMFTTGLYTCGRFVLMFKFDEFTFLKTDYKCSVILVPTLPAIINSEL 300
 QY 303 LDKYDLSNLVEIASGAPLAKSKEIGAVARFNLPVROGVGLTETTSALITTPGDDPKG 362
 |||||
 Db 301 VDKYDLSNLVEIASGAPLAKSKEIGAVARFNLPVROGVGLTETTSALITTPGDDPKG 360
 QY 363 ASGVKVPLEFAKAVYIDLTFTKTLGPNRRGEVCKGPMILKKGVDNPEATREIIDEESWLT 422
 |||||
 Db 361 AGCKVVPLEFAKAVYIDLTFTKTLGPNRRGEVCKGPMILKKGVDNPEATREIIDEESWLT 420
 QY 423 GDIGYDEEHFFIVDLRLSLIKYGYVPPALESVLLQHPNIDAGVAGVDPPIAGEL 482
 |||||
 Db 421 GDIAVYDKDGHFFIVDLRLSLIKYGYVPPALESVLLQHPNIDAGVAGVDPPIAGEL 480
 QY 483 PGAVVVLKKGKSMTEKEMDVYASQVSNARKLRGVRFEVDEVRKGLTGKIDGKAIREIL 541
 |||||
 Db 481 PAAVVVLKKGKSMTEKEMDVYASQVSNARKLRGVRFEVDEVRKGLTGKIDGKAIREIL 539

RESULT 5
 A26772
 Photinus-luciferin 4-monoxygenase (ATP-hydrolyzing) (EC 1.13.12.7) - common eastern
 N:Alternate names: firefly luciferase
 C:Species: Photinus pyralis (common eastern firefly)
 C>Date: 19-Nov-1988 #sequence_revision 20-Aug-1994 #text_change 26-May-2000
 C:Accession: A26772
 R:de Wet, J.R.; Wood, K.V.; Deluca, M.; Hellinski, D.R.; Subramani, S.
 Mol. Cell. Biol. 7, 725-737, 1987
 A:Title: Firefly luciferase gene: structure and expression in mammalian cells.
 A:Reference number: A26772; MUID:87144243
 A:Accession: A26772
 A:Molecule type: DNA
 A:Residues: 1-550 <DEM>
 A:Cross-references: GB:M15077; NID:g160793; PIDN:AAA29795.1; PID:g160794
 A:Note: The authors translated the codon CAA for residue 134 as Glu
 C:Comment: This protein catalyzes the oxidation of luciferin in the presence of ATP,
 species from green to yellow.
 C:Genetics:
 A:introns: 42/3; 112/1; 223/2; 330/2; 448/3; 504/1
 C:Superfamily: 4-comumate--CoA ligase; acetate--CoA ligase homology
 C:Keywords: ATP; luminescence; monoxygenase; oxidoreductase
 F:71-535/Domain: acetate--CoA ligase homology <ACL>
 F:548-550/Region: peroxisome/glyoxysome location signal (S-[RKH]-L) motif

Query Match 68.9%; Score 1945.5; DB 1; Length 550;
 Best Local Similarity 68.2%; Pred. No. 1.5e-130;
 Matches 369; Conservative 75; Mismatches 96; Indels 1; Gaps 1;

QY 4 MENDENIVYGEPEPYPIEESAGAOQLRKRYMDYAKL-GATAFTNALTGVDYTAELYEKS 62
 |||||
 Db 1 MEDAKNIHKGAPFYPLEDEGTAGEOLHKAMKRYAGVGTIAFTNAHEVNTIYEYEMA 60
 QY 63 CCLGEALKNYGLVVDGRIALCSECEFFIPVLAGLFTGVGAPATNEITYTLRELVHSLGI 122
 |||||
 Db 61 VRLAEAMKRYGLTNRHIVCSENSLOPFMPVCGALFVGAVAPADIDYNEBELLSMNI 120
 QY 123 SKPTIVSSKGLDKVITVQVTAIKTIYVLDKVDYRGQSDNFIKKNTPGFGSS 182
 |||||
 Db 121 SOPTIVSSKGLDKVITVQVTAIKTIYVLDKVDYRGQSDNFIKKNTPGFGSS 180
 QY 183 FKTVEVNRKDOVALIMNMSGSTGLPKGVOLTHENIVTRFSHARDPIYGNQVSPGAILTV 242
 |||||
 Db 181 FVPEFDRDKTIALIMNMSGSTGLPKGVOLTHENIVTRFSHARDPIYGNQVSPGAILTV 240
 QY 243 VPFHNGFMFTTGLYTCGRFVLMFKFDEFTFLKTDYKCSVILVPTLPAIINSEL 302
 |||||
 Db 241 VPFHNGFMFTTGLYTCGRFVLMFKFDEFTFLKTDYKCSVILVPTLPAIINSEL 300
 QY 303 LDKYDLSNLVEIASGAPLAKSKEIGAVARFNLPVROGVGLTETTSALITTPGDDPKG 362
 |||||
 Db 301 IDKYDLSNLVEIASGAPLAKSKEIGAVARFNLPVROGVGLTETTSALITTPGDDPKG 360
 QY 363 ASGVKVPLEFAKAVYIDLTFTKTLGPNRRGEVCKGPMILKKGVDNPEATREIIDEESWLT 422
 |||||
 Db 361 AVGVKVPLEFAKAVYIDLTFTKTLGPNRRGEVCKGPMILKKGVDNPEATREIIDEESWLT 420
 QY 423 GDIGYDEEHFFIVDLRLSLIKYGYVPPALESVLLQHPNIDAGVAGVDPPIAGEL 482
 |||||
 Db 421 GDIAVYDEEHFFIVDLRLSLIKYGYVPPALESVLLQHPNIDAGVAGVDPPIAGEL 480
 QY 483 PGAVVVLKKGKSMTEKEMDVYASQVSNARKLRGVRFEVDEVRKGLTGKIDGKAIREIL 542
 |||||
 Db 481 PAAVVVLKKGKSMTEKEMDVYASQVSNARKLRGVRFEVDEVRKGLTGKIDGKAIREIL 540
 QY 543 K 543
 |||||
 Db 541 K 541

RESULT 6

S29354

Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) YE [validated] - Jun

N:Alternate names: firefly-type luciferase

C:Species: Pyrophorus plagiophthalmus

C:Date: 22-Nov-1993 #sequence_revision 01-Sep-1995 #text_change 02-Jun-2000

C:Accession: S29354

R:Wood, K.V.; Lam, Y.A.; Seliger, H.H.; McElroy, W.D.

Science 244, 700-702, 1989

A:Title: Complementary DNA coding click beetle luciferases can elicit bioluminescence of

A:Reference number: S29352; MUID:89242142

A:Accession: S29354

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-543 <WOO>

C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology

C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase

F:70-532/Domain: acetate--CoA ligase homology <ACL>

Query Match 48.4%; Score 1367; DB 2; Length 543;

Best Local Similarity 48.3%; Pred. No. 2e-89;

Matches 263; Conservative 108; Mismatches 167; Indels 6; Gaps 3;

```

OY 4 MENDNIYVGPPEFYPIEGSAGAOLRKMDRYAKLGAIAFTNALTGVDTTAEYLEKSC 63
DB 2 MKREKNVIYGPPEPLHLEDTAGEMLFRLRKSHLPQ-ALVDVFGDESLSYKEFEATC 60
OY 64 CLGELKNYGLVVDGRIALCSECEFFIPVLAGLFIYGAVAPNTETTLRELHSLGIS 123
DB 61 LLAQSLHNGCYKMDVVSICAENNRFFPIIAWYIGMIVAPVNSYIPDELCKVMGIS 120
OY 124 KPTVFSKKGLDKVITVQKTYTAIKTYIILDSKVYRGQSDMNFIRKNTPOGFGSSF 183
DB 121 KPQIVCTKNTLTKVLEVOGRTFIKRIIILDTVENIHGCESLPNFISYSDGNT--ANF 178
OY 184 KTVENRKEQVALINSSGSTGLPKGVOLTHENITRSHADPIYGNVSPGTALIVY 243
DB 179 KPLHNDPVEQVAAILCSTGTGLPKGVQTHONICRLIHADPAGTOLIPGVVLVYV 238
OY 244 PFHGFMTTGLYLTGCFRIYMLTRFDETFKLTDYKCSSVILVPLFAILRSELL 303
DB 239 PFHAFGFSINIGYPMVGLRVIMLRFOEAPLKAIDYEVASVINVPAIILFLSKSPLY 298
OY 304 DKYDLSNIVELASGAPLSKEIGEAVARFNLPGVROGGLTETTSAILITPEGDCKPCA 363
DB 299 DKYDLSNIVELASGAPLSKEIGEAVARFNLPGVROGGLTETTSAILITPEGDCKPCA 363
OY 364 SGKVVPLFKAKVIDIDTKTGLPNRGEVCYKPMIMGYNDPNPATREIIDESEMLTNG 423
DB 359 LGRVPLMAAKIADRETKALGPNOVGELCIKGPVNSGYNNVNAVATKEAIDDDGMLHSG 418
OY 424 DIGYDEKHEFFIYDRKSLIKYKGYVPPALESVLLQHPNIFDAGVAGVDPPIAGELP 483
DB 419 DFGYDEDEHFFVYDRKSLIKYKGYVPPALESVLLQHPNIFDAGVAGVDPPIAGELP 478
OY 484 GAVVVLKKGSMTEKEVNDYASOVSNARKRLRGVRFVDEVPKGLTGKIDKALREILKK 543
DB 479 SAFFVLDGKREITAKKEYDYLAERVSHTKYLRGVRFDVSPRNVTKITRK--ELLKQ 535
OY 544 PVAK 547
DB 536 LLEK 539

```

RESULT 7

Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) YG [validated] - Jun

N:Alternate names: firefly-type luciferase

C:Species: Pyrophorus plagiophthalmus

C:Date: 25-Feb-1994 #sequence_revision 01-Sep-1995 #text_change 02-Jun-2000

C:Accession: S29353

R:Wood, K.V.; Lam, Y.A.; Seliger, H.H.; McElroy, W.D.

Science 244, 700-702, 1989

A:Title: Complementary DNA coding click beetle luciferases can elicit bioluminescence

A:Reference number: S29352; MUID:89242142

A:Accession: S29353

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-543 <WOO>

C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology

C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase

F:70-532/Domain: acetate--CoA ligase homology <ACL>

Query Match 48.4%; Score 1367; DB 2; Length 543;

Best Local Similarity 49.1%; Pred. No. 2e-89;

Matches 267; Conservative 103; Mismatches 168; Indels 6; Gaps 3;

```

OY 4 MENDNIYVGPPEFYPIEGSAGAOLRKMDRYAKLGAIAFTNALTGVDTTAEYLEKSC 63
DB 2 MKREKNVIYGPPEPLHLEDTAGEMLFRLRKSHLPQ-ALVDVFGDESLSYKEFEATC 60
OY 64 CLGELKNYGLVVDGRIALCSECEFFIPVLAGLFIYGAVAPNTETTLRELHSLGIS 123
DB 61 LLAQSLHNGCYKMDVVSICAENNRFFPIIAWYIGMIVAPVNSYIPDELCKVMGIS 120
OY 124 KPTVFSKKGLDKVITVQKTYTAIKTYIILDSKVYRGQSDMNFIRKNTPOGFGSSF 183
DB 121 KPQIVCTKNTLTKVLEVOGRTFIKRIIILDTVENIHGCESLPNFISYSDGNT--ANF 178
OY 184 KTVENRKEQVALINSSGSTGLPKGVOLTHENITRSHADPIYGNVSPGTALIVY 243
DB 179 KPLHNDPVEQVAAILCSTGTGLPKGVQTHONICRLIHADPAGTOLIPGVVLVYV 238
OY 244 PFHGFMTTGLYLTGCFRIYMLTRFDETFKLTDYKCSSVILVPLFAILRSELL 303
DB 239 PFHAFGFSINIGYPMVGLRVIMLRFOEAPLKAIDYEVASVINVPAIILFLSKSPLY 298
OY 304 DKYDLSNIVELASGAPLSKEIGEAVARFNLPGVROGGLTETTSAILITPEGDCKPCA 363
DB 299 DKYDLSNIVELASGAPLSKEIGEAVARFNLPGVROGGLTETTSAILITPEGDCKPCA 363
OY 364 SGKVVPLFKAKVIDIDTKTGLPNRGEVCYKPMIMGYNDPNPATREIIDESEMLTNG 423
DB 359 LGRVPLMAAKIADRETKALGPNOVGELCIKGPVNSGYNNVNAVATKEAIDDDGMLHSG 418
OY 424 DIGYDEKHEFFIYDRKSLIKYKGYVPPALESVLLQHPNIFDAGVAGVDPPIAGELP 483
DB 419 DFGYDEDEHFFVYDRKSLIKYKGYVPPALESVLLQHPNIFDAGVAGVDPPIAGELP 478
OY 484 GAVVVLKKGSMTEKEVNDYASOVSNARKRLRGVRFVDEVPKGLTGKIDKALREILKK 543
DB 479 SAFFVLDGKREITAKKEYDYLAERVSHTKYLRGVRFDVSPRNVTKITRK--ELLKQ 535
OY 544 PVAK 547
DB 536 LLEK 539

```

RESULT 8

Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) OR [validated] -

N:Alternate names: firefly-type luciferase

C:Species: Pyrophorus plagiophthalmus

C:Date: 22-Nov-1993 #sequence_revision 03-Nov-1995 #text_change 02-Jun-2000

C:Accession: S29355

R:Wood, K.V.; Lam, Y.A.; Seliger, H.H.; McElroy, W.D.

Science 244, 700-702, 1989

A:Title: Complementary DNA coding click beetle luciferases can elicit bioluminescence

A:Reference number: S29352; MUID:89242142

A:Accession: S29355

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-543 <WOO>

C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology

C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase


```

QY 163 YOSMDFIKKNTPOGFKSSFKTEVNRKROYALIMNSSGTLGPKVOLTHENTYTR-- 220
D 163 FSKLMADSEME-----VIVNSDDVAL-PRSSGTLGLPGVMLTHKGLVTSVA 212
QY 221 -----FSHARDPIYGNVSPGTALIVVPHHGFEMTTLGLYTCGFR---IYM 266
D 213 QOVDGDNPMIYHSESD-----VMICILPLFH---IYSLNAVLCGGIRAGVTTILI 258
QY 267 LKRFDEETFLKTLQDYKSSVILVPTFLAIRSELIDKVDLSNVEIASGAPLSKEIG 326
D 259 MOKFDIVPELELOKTKVTITGPVPIYALAKSPVDYDLSVTVMSGAAPLGKELE 318
QY 327 EAVARRFNLPGVRQGYGLTETTSATII-----TPRGDDKPGASGVVPLFKAKVIDLDTK 381
D 319 DAVRAKFPNAKIGQGYGTEAGPVLAMCLAFKPEYIKSGACGTVVNAEMKIVDPETN 378
QY 382 KTLGPNRGEVCGVGMIMKGVNDPEATREIIDEEGMLHTGDIYDEEKFHFIVDRK 441
D 379 ASLPRNKGELICIGDQIMKGYLNDPESTRITIDEEGMLHTGDIIDDELFIYDRK 438
QY 442 SLIKYGYVPAPELESVLLQHPNIFDAGVAGVDPPIAGELPGAVVVLKCKGSMTEKEVM 501
D 439 EIIKKGQVAPAELEALLLTHPTISDAVAVPMIDEKAGEVPVAVVATNGFTTEBEIK 498
QY 502 DYVASQVSNARKLGVNRPVDEVPKGLTGKIDKKAIR 538
D 499 QFVSKQVVFYKRI-FRVFVDAIPKSPSGKILRKDLR 534

```

RESULT 11

```

S15695
4-comumate--CoA ligase (EC 6.2.1.12) (clone Pc4CL-2) - parsley
C:Species: Petroselinum crispum (parsley)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-May-2000
C:Accession: S15695
R:Lozova, E.; Hoffmann, H.; Douglass, C.; Schulz, W.; Scheel, D.; Hahlbrock, K.
Eur. J. Biochem. 176, 661-667, 1988
A:Title: Primary structures and catalytic properties of isoenzymes encoded by the two 4-
A:Reference number: S01667; MUID:89005119
A:Accession: S15695
A:Molecule type: mRNA
A:Residues: 1-544 <L02>
A:Cross-references: EMBL:X13325; NID:920435; PIDN:CAA1697.1; PID:920436
A:Gene: 4CL-2
C:Genetics:
A:Superfamily: 4-comumate--CoA ligase; acetate--CoA ligase homology
C:Keywords: acid-thiol ligase; coenzyme A
F:74-533/Domain: acetate--CoA ligase homology <ACLD>

```

```

Query Match 26.9%; Score 758.5; DB 2; Length 544;
Best local Similarity 35.8%; Pred. No. 3.6e-46;
Matches 185; Conservative 89; Mismatches 190; Indels 53; Gaps 10;
QY 46 NALTGVDTYAAVLEKSCCLGALKNYGLVNDGRIALSCENCEFFIPVLAGLFGVGA 105
D 47 NGATGEFTYSQVLESLSRKVASGLNKIGQGTIMLLNSPEYFAFAGASTRGALST 106
QY 106 PTNEITLRELVLHSLGSKPTIVFSSKKGIDKYTVOKTVAIKTIYIL--DSKYDYG 162
D 107 MANPFTSAEVIKQLKASLKLITQACVYDK---KDAAEKNIOIICIDAPQCLH 162
QY 163 YOSMDFIKKNTPOGFKSSFKTEVNRKROYALIMNSSGTLGPKVOLTHENTYTR-- 220
D 163 FSKLMADSEME-----VIVNSDDVAL-PRSSGTLGLPGVMLTHKGLVTSVA 212
QY 221 -----FSHARDPIYGNVSPGTALIVVPHHGFEMTTLGLYTCGFR---IYM 266
D 213 QOVDGDNPMIYHSESD-----VMICILPLFH---IYSLNAVLCGGIRAGVTTILI 258
QY 267 LKRFDEETFLKTLQDYKSSVILVPTFLAIRSELIDKVDLSNVEIASGAPLSKEIG 326

```

```

D 259 MOKFDIVPELELOKTKVTITGPVPIYALAKSPVDYDLSVTVMSGAAPLGKELE 318
QY 327 EAVARRFNLPGVRQGYGLTETTSATII-----TPRGDDKPGASGVVPLFKAKVIDLDTK 381
D 319 DAVRAKFPNAKIGQGYGTEAGPVLAMCLAFKPEYIKSGACGTVVNAEMKIVDPETN 378
QY 382 KTLGPNRGEVCGVGMIMKGVNDPEATREIIDEEGMLHTGDIYDEEKFHFIVDRK 441
D 379 ASLPRNKGELICIGDQIMKGYLNDPESTRITIDEEGMLHTGDIIDDELFIYDRK 438
QY 442 SLIKYGYVPAPELESVLLQHPNIFDAGVAGVDPPIAGELPGAVVVLKCKGSMTEKEVM 501
D 439 EIIKKGQVAPAELEALLLTHPTISDAVAVPMIDEKAGEVPVAVVATNGFTTEBEIK 498
QY 502 DYVASQVSNARKLGVNRPVDEVPKGLTGKIDKKAIR 538
D 499 QFVSKQVVFYKRI-FRVFVDAIPKSPSGKILRKDLR 534

```

RESULT 12

```

T07909
4-comumate--CoA ligase (EC 6.2.1.12) 1 [validated] - western balsam poplar x cottonw
C:Species: Populus trichocarpa x Populus deltoides (western balsam poplar x cottonw
C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 26-May-2000
C:Accession: T07909
R:Allina, S.M.; Pri-Hadash, A.; Theilmann, D.A.; Ellis, B.E.; Douglas, C.J.
Plant Physiol. 116, 743-754, 1998
A:Title: 4-Comumate:Coenzyme A ligase in hybrid poplar. Properties of native enzymes
A:Reference number: Z16208; MUID:98150279
A:Accession: T07909
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-557 <ALD>
A:Cross-references: EMBL:AF008184; NID:92911798; PIDN:AAC39366.1; PID:92911799
A:Gene: 4CL1
C:Function:
A:Description: EC 6.2.1.12; provides activated thioester substrates for phenylpropano
C:Superfamily: 4-comumate--CoA ligase; acetate--CoA ligase homology
C:Keywords: acid-thiol ligase; coenzyme A; flavonoid biosynthesis
F:74-534/Domain: acetate--CoA ligase homology <ACLD>

```

```

Query Match 26.8%; Score 756; DB 2; Length 557;
Best local Similarity 35.5%; Pred. No. 5.9e-46;
Matches 183; Conservative 84; Mismatches 199; Indels 50; Gaps 11;
QY 46 NALTGVDTYAAVLEKSCCLGALKNYGLVNDGRIALSCENCEFFIPVLAGLFGVGA 105
D 47 NGFTGIIHTVAEVELTSRKVASGLNKIGQGDVILLILONSPEYFAFAGASTIIGAI 106
QY 106 PTNEITLRELVLHSLGSKPTIVFSSKKGIDKYTVOKTVAIKTIYILDSVNDYRGVS 165
D 107 TANPFTPAEVAKQATASAKLITQAVYAEKVQFENHVKIYIVDSPEENTLHSE 166
QY 166 MOKFIKNTPOGFKSSFKTEVNRKROYALIMNSSGTLGPKVOLTHENTYTR-- 220
D 167 LTNSEDDIP-----AVEINPDVVAL-PRSSGTLGLPGVMLTHKGLVTSVAQOV 216
QY 221 -----FSHARDPIYGNVSPGTALIVVPHHGFEMTTLGLYTCGFR---IYMLTK 269
D 217 DQENRILYHEHD-----VILCVLPLFH---IYSLNAVLCGLRGSAILLMOK 262
QY 270 FDEETFLKTLQDYKSSVILVPTFLAIRSELIDKVDLSNVEIASGAPLSKEIGAV 329
D 263 FEIVTLMELVQKRYKTIAPFVPPVLAIVAKCPYDKDYLSSIRYMSGAAPMGELEBY 322
QY 330 ANRFLNPGVR--QGYGLTET--TSALITITPEG--DOKPGASGVVPLFKAKVIDLDTK 382
D 323 --RAKLPNAKIGQGYGTEAGPVLAMCLAFKPEYIKSGACGTVVNAEMKIVDPETN 380
QY 383 TLGPNRGEVCGVGMIMKGVNDPEATREIIDEEGMLHTGDIYDEEKFHFIVDRK 442

```

Db 381 SLPRNDSGEICIRGSGIMKGYLNDEPATERIVDNDGMLHTTGIDIGYIDGDELFIVDRKE 440
 QY 443 LKRYGYOVPAPLESVLLQHNIFPDAGVAGVDDPIAGCLPGAVVVLKKGKSMTEKEVMD 502
 Db 441 LKRYGFOVAPAELEMLAIHPDISCAVAPKMDKDEAGVPIAFVVRANGSKITEDEIKQ 500
 QY 503 YVASOVNAKRLRGVRFVDEVPKGLTGKIDKPAIR 538
 Db 501 YISKOVIFYKRI-SRVEFFTEALPKAPSGKILKDKLR 535

RESULT 13

H85064

4-coumarate--CoA ligase-like protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001
 C:Accession: H85064
 R:anonymus, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
 Nature 402, 769-777, 1999
 A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A:Reference number: AB5001; MUID:20083488

A:Accession: H85064
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-544 <STO>
 A:Cross-references: GB:NC_001268; NID:g7267275; PIDN:CAB81058.1; GSPDB:GN00140
 C:Genetics:
 A:gene: AT4G05160
 A:Map position: 4
 C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology

Query Match 26.7%; Score 754; DB 2; Length 544;
 Best Local Similarity 34.0%; Pred. No. 7.9e-46;
 Matches 178; Conservative 111; Mismatches 195; Indels 40; Gaps 12;

QY 42 IAFNTALGVDTYAEYLEKSCCLGALKNGVLDGRIALCSENCPEFFIPVLAGLFTG 101
 Db 44 LATADSDTGLTFSSQLKSAVARLHGFRLGIRKNDVLIAPNSYQPLCEFLAVTAIG 103
 QY 102 VGAPAPNEIYTLRELHVSIGISKPTIV-----FSSKGLDKVITYQKIVTAIKTIIVLD 155
 Db 104 GVFETNAPLYTAVNEVSKQIKDSNPKTIISVNOLEFKIKFEDLPVVLGSKDVEIPEPGSN 163
 QY 156 SKVDYAGYOSMDNFIKKNPQOGKSSFKTVEYNRKEQVALIMNSSGSGTGPVOLTHE 215
 Db 164 SKI-----LSFDNVMELSEPV-----SEYFPEVKQSDTALLY-SSGTTGTSKVEELTHG 213
 QY 216 N-----IVTRESHADPIYGNQVSPGTAIILVPPFHGFGM-FTTLGYLTCGFRIVMLT 268
 Db 214 NFIAASLWMTMODDLNGEYHG-----VELCFLPWFHVFGLAVITYSQLOGNMLVSMA 266
 QY 269 KPEEEFLTKLDYKCSSVILVPTLEAILNRSLDKYDLSNIVETASGAPLSKETGA 328
 Db 267 RFELEVLVKNIEKFRVTHLMVPPVFLALSKQISYKFKPLDSLKYSGSAPLGLKMLBE 326
 QY 329 VAREFNLPGV--RQGYGLTETTSAILITPE--GDOPKSGAKVPLPFKAKVIDLDTKTL 384
 Db 327 CGR--NIPNVLMOGYGMETGCIIVSEDPRLGKRSSGAGMLAPVEVAQIYSEVDEKSG 384
 QY 385 GPNRREGVCYKGMKKGIVDNEATRETIIDEGMLHTGDIGYDEKHFPIVDRLSLI 444
 Db 385 PPNQOGEIIVWKGPMKGYLNINPQATKETIDKSKSWHTGDLGTFNEDGMLVYVDRIKELI 444
 QY 445 KYKGYOVPAPELESVLLQHNIFPDAGVAGVDDPIAGCLPGAVVVLKKGKSMTEKEVMDV 504
 Db 445 KYKGYOVPAPELESVLLQHNIFPDAGVAGVDDPIAGCLPGAVVVLKKGKSMTEKEVMDV 504
 QY 505 ASOVNAKRLRGVRFVDEVPKGLTGKIDKPAIR 548
 Db 505 AKQVAPYKRLR-RVPSFLSLVPRKSAAKI---LRRELVOQVRSKM 544

RESULT 14

T07908

4-coumarate--CoA ligase (EC 6.2.1.12) 2 - western balsam poplar x cottonwood
 C:Species: Populus trichocarpa x Populus deltoides (western balsam poplar x cottonwood)
 C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 11-May-2000
 C:Accession: T07908
 R:Allina, S.M.; Pri-Hadash, A.; Theilmann, D.A.; Ellis, B.E.; Douglas, C.J.
 Plant Physiol. 116, 743-754, 1998
 A:Title: 4-Coumarate:Coenzyme A ligase in hybrid poplar. Properties of native enzymes
 A:Reference number: Z16208; MUID:98150279
 A:Accession: T07908
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-548 <ATL>
 A:Cross-references: EMBL:AF008183; NID:g2911796; PIDN:AC039365.1; PID:g2911797
 C:Genetics:
 A:gene: 4CL2
 C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology
 C:Keywords: acid-thiol ligase; coenzyme A; flavonoid biosynthesis
 F:74-533/Domain: acetate--CoA ligase homology <ACL>

Query Match 26.7%; Score 753.5; DB 2; Length 548;
 Best Local Similarity 36.0%; Pred. No. 8.6e-46;
 Matches 186; Conservative 86; Mismatches 193; Indels 51; Gaps 12;

QY 46 NALTGVDYTAEYLEKSCCLGALKNGVLDGRIALCSENCPEFFIPVLAGLFTGVGA 105
 Db 47 NGPTDITTVADVLTSSKVASGLKGLGQGDVILLILNOSPFEVAFAGASTGAISS 106
 QY 106 PTNEIYTLRELHVSIGISKPTIVSSKKGLDITVQKIVTAIKTIIVILDSKVDYRGYS 165
 Db 107 TANPPTSAELAKQATASKAKLIITHAAVAKVQOQFQENDHVK-IMTIDSLT----- 158
 QY 166 MDNFIRKNTPOGKSSKTYEVNRKEQVALIMNSSGSGTGPVOLTHEIVYR----- 220
 Db 159 -ENCLHFSLELTSSDENELPTVKI- KPDIDMALPYSSGTTGLPKGVMLTHKGLTVSAQV 216
 QY 221 -----FSHADPPIYGNQVSPGTAIILVPPFHGFGMFTTLGYLTCGFR---IVMLTK 269
 Db 217 DGENNLYFHERD-----VILCVLPFH---IYLSNVSFLGLRAGSAILVMQK 262
 QY 270 FDETFELTKLDYKCSSVILVPTLEAILNRSELDDKYDLSNIVETASGAPLSKETGEAV 329
 Db 263 FDTVALMDLVQYKKTIVLAPVLPICLAIKSPVVDQYDLSIRIVYLSGAPLGELEDTV 322
 QY 330 ARFNLPGVR--QSGILETTSAILI-----TPESDDKPGASGVPLPFKAKVIDLDTK 382
 Db 323 --RAKLPNAKLGQGYGMTEAGPVIAMCLAFKPEPEIKSGAGCTVYRNAEKIYDPEPGE 380
 QY 383 TLGPNRREGVCYKGMKKGIVDNEATRETIIDEGMLHTGDIGYDEKHFPIVDRLS 442
 Db 381 SOPRKRKTGECIRGCQIMKGYLNINPQATKETIDKQWHTGDLGIDIDE-LFIVDRKE 439
 QY 443 LKRYGYOVPAPLESVLLQHNIFPDAGVAGVDDPIAGCLPGAVVVLKKGKSMTEKEVMD 502
 Db 440 LKRYGFOVAPAELEMLAIHPNISDAVVPKMDKDEAGVPAFVVRANGSKITEDEIKQ 499
 QY 503 YVASOVNAKRLRGVRFVDEVPKGLTGKIDKPAIR 538
 Db 500 YISKOVIFYKRI-GRVFFTEALPKAPSGKILKDKLR 534

RESULT 15

T03789

4-coumarate--CoA ligase (EC 6.2.1.12) 4CL2 - common tobacco
 N:Alternate names: 4-coumaroyl-CoA synthetase
 C:Species: Nicotiana tabacum (common tobacco)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
 C:Accession: T03789
 R:Lee, D.; Douglas, C.J.
 Plant Physiol. 112, 193-205, 1996
 A:Title: Two divergent members of a tobacco 4-coumarate:coenzyme A ligase (4CL) gene

A:Reference number: 215086; MUID:96416441

A:Accession: T03789

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1542 <LEF>

A:Cross-references: EMBL:U50846; NID:g1663723; PIDN:AAB18638.1; PID:g1663724

C:Genetics:

A:Gene: ACL2

C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology

C:Keywords: acid-thiol ligase; coenzyme A; flavonoid biosynthesis

F:73-53/Domain: acetate--CoA ligase homology <ACL>

Query Match 26.4%; Score 746.5; DB 2; Length 542;

Best Local Similarity 34.9%; Pred. No. 2,7e-45;

Matches 189; Conservative 90; Mismatches 200; Indels 63; Gaps 12;

```
QY 32 YMDRAKAGIAFTN-----ALTVD---YTYAELEKSCCIGBALKNYGLVYDGR 79
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 20 YIPNHLPLHSYCFENISESSRPLINGANKQIYADVDELNRKVAAGLHKQGIQPKDT 79
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 80 IALCENCEEFPIPIAGLFIQVVAAPTNEIYTLRELVLHSLGISKPTIVFSSKKGLDKVI 139
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 80 IMILLPNSPEFPAFTGASYLCAISTMANPLTPAIVAKQAKASSAKITVTOACHVNVK- 138
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 140 TVQKTVTAKITVILDSKVYRGYSMDNFIK---KNTPOGFGSSFKT-----V 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 139 -----KDY---AFENDVKIICIDSAPEGCLHFSVLTQANENHDIPV 176
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 187 EYNRKEQVALINSSGSLPFGVQTLHENITVRESHARDPIYGN-QVSPGTAILTVPF 245
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 177 ELQPDVVAL-PPSSGTTGLPKGMVTHKGLVTSVAQOVGPNMLYIHSEDMCLVPL 235
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 246 HHGFMFTLIGLTGCFR---IVMLTKFEETPLKTLQDYKCSSYILVPTLFALINRSE 301
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 236 FH---IYGLNSVLLGSLVGAAILIMQKFDIYFELIORYKVTIGPEVPIVLAIAKSP 292
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 302 LIDKYDLSNLVEIASGAPLSKEIGEAVARFPNLPVGROGYGLTETSAIII-----TPE 356
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 293 MVDYDLSVFTVMSGAAPLGKELKEDTVAKFPNNAKLGCGYGMTAGPYLAMCLAFAPKEP 352
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 357 GDDKPGASGKVPPLTKAKVIDLDTKTLGPNRGEVCVCPMLMKGYVNDPEATREITIDE 416
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 353 FEIKSGACGTYVRNAEMKIVDPKTSNLSLRNOSGEICIRGQIMKGYLNDPEATARTIDK 412
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 417 EGMVLTGDIQYDEKHEFIYDRLKSLIKYKGQYQVPAELESVLLQHPNIFDAGVAGVPD 476
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 413 EGMVLTGDIQYIDDDDELFIYDRLKELIKYKGQYQVPAELESVLLQHPNISDAAVVPMKD 472
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 477 PIAGELPGAVVYLKKGKSMTEKEVMQVYASQVSNAKRLRGVRFVDEVKGLTGKIDGKA 536
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 473 E0AGEVPAFVVRNSGSTITDEVDKDFISKQVLYFYKRIK-RVFEVDALPKSPSGILRKD 531
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 537 IR 538
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 532 LR 533
```

Search completed: November 29, 2001, 01:57:52
Job time: 362 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2001, 01:53:15 ; Search time 64.96 Seconds
(without alignments)
1233.948 Million cell updates/sec

Title: US-09-581-241-4
Perfect score: 2823
Sequence: 1 MEMMENDENIYGPPEPTPI.....TGKIDGKAIRELKKPVAKM 548

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL.17:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2793	98.9	548	5	027321 luciola lat
2	2792	98.9	548	5	027348 luciola lat
3	2343.5	83.0	548	5	025118 hotaria par
4	2335.5	82.7	548	5	026304 luciola min
5	1965.5	69.6	547	5	027688 lampyris no
6	1941.5	68.8	550	5	027755 photinus py
7	1918.5	68.0	550	5	027758 photinus py
8	1898	67.2	548	5	026076 pyrococella
9	1890	67.0	548	5	026076 pyrococella
10	1831	64.9	552	5	094697 photuris pe
11	1825	64.6	552	5	094696 photuris pe
12	1623.5	57.5	545	5	027757 photuris pe
13	1459.5	51.7	545	5	094408 phrixothrix
14	1330	47.1	546	5	094407 phrixothrix
15	1025	36.3	544	5	09VCC6 drosophila
16	775.5	27.5	546	10	09VCC6 drosophila
17	773	27.4	636	10	042879 lithospermu
18	772	27.3	562	10	09FGM4 arabidopsis
19	758	26.9	557	10	048869 populus dal

20	755.5	26.8	548	10	048868	048868 populus dal
21	755	26.7	544	10	09M0X9	09M0X9 arabidopsis
22	748.5	26.5	580	10	09LMV8	09LMV8 arabidopsis
23	748	26.5	544	10	09LL50	09LL50 rubius idae
24	744	26.4	543	10	09LL49	09LL49 rubius idae
25	744	26.4	585	10	042880	042880 lithospermu
26	740.5	26.2	542	10	042943	042943 nicotiana t
27	735.5	26.1	597	5	09Y158	09Y158 drosophila
28	734.5	26.0	535	10	081139	081139 populus tre
29	734.5	26.0	545	10	08SMW8	08SMW8 solanum tub
30	721	25.5	1549	10	09LMV7	09LMV7 arabidopsis
31	720	25.5	544	5	019339	019339 caenorhabdi
32	715	25.3	570	10	081140	081140 populus tre
33	712	25.2	522	2	09K3W1	09K3W1 streptomyc
34	693	24.5	556	10	09M7S2	09M7S2 lolium pere
35	684	24.2	570	10	09LUJ6	09LUJ6 arabidopsis
36	676.5	24.0	542	10	09LQ12	09LQ12 arabidopsis
37	675.5	23.9	570	10	09M7S3	09M7S3 lolium pere
38	674	23.9	591	10	09LL51	09LL51 rubius idae
39	666.5	23.6	542	10	09FOX7	09FOX7 capsicum an
40	657	23.3	557	10	09M7S1	09M7S1 lolium pere
41	649	23.0	544	10	09FP44	09FP44 arabidopsis
42	635	22.5	532	1	028762	028762 archaeoglob
43	634.5	22.5	566	10	049414	049414 arabidopsis
44	602	21.3	566	2	09K8A2	09K8A2 bacillus ba
45	599.5	21.2	569	1	028502	028502 archaeoglob

ALIGNMENTS

RESULT 1
ID 027321 PRELIMINARY; PRT; 548 AA.
AC 027321;
DT 01-NOV-1996 (TREMUREL. 01, Created)
DT 01-NOV-1996 (TREMUREL. 01, Last sequence update)
DT 01-JUN-2001 (TREMUREL. 17, Last annotation update)
DE LUCIFERASE.
OS Luciola lateralis (Firefly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Elateriformia; Cantharoidae; Lampyridae; Luciola.
OX NCBI_TaxID=7052;
RN [1]
RP SEQUENCE FROM N.A.
RA Cho K., Choi Y., Boo K.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=MJUU; TISSUE=ABDOMEN;
RC Cho K.H., Choi Y., Boo K.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN=MJUU; TISSUE=ABDOMEN;
RC Cho K., Choi Y., Boo K.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z69619; CAA93444.1; -;
DR EMBL: U49181; AAA91471.1; -;
DR EMBL: U51019; AAB00229.1; -;
DR HSSP: P08659; ILCT.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00455; AMP-BINDING; 1.
SQ SEQUENCE 548 AA: 60048 MW; E26C6C26F423E5E8 CRC64;

Query Match 98.9%; Score 2793; DB 5; Length 548;
Best Local Similarity 98.7%; Pred. No. 5.6e-185;
Matches 541; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 MEMMENDENIYGPPEPTPIEGSGAGLRLKRYMRYKALGAIATPTNLVTGVDTYAYYLE 60

Query Match 98.9%; Score 2792; DB 5; Length 548;

Query Match	83.0%;	Score 2343.5;	DB 5;	Length 548;
-------------	--------	---------------	-------	-------------

Best Local Similarity 81.98; Pred. NO. 7e-154;
Matches 447; Conservative 47; Mismatches 51; Indels 1; Gaps 1.

QY	4	MENDEINVPPEPFPPPIEESGAGQJRKVYDRYAKGAIATFALGVYATVAYEIKSG	63
Db	3	MEKENVNVGPRLPFYPIEESAGIQJHKWQYAKIGAIATFALGVDSYGEYEDTIC	62
QY	64	CLGELAKNTGLVVDGRJALCSNCEFFPIYVLAGLPIGVAVPTNETYITRELVHSLGIS	123
Db	63	RLAEAMNVMQKCEGTIALCSNCEFFPIYVLAGLIGAVAVPTNETYITRELHNSHSLGIA	122
QY	124	KPTIVFSSKKGIDKVIIVOKTVAIKTIVILDSKVDRGYOSMDNFJTKKTPGCFKGSF	183
Db	123	QPTIVFSSKKGIDKVIIVOKTVAIKTIVILDSKVDRGYOSMDNFJTKKTPGCFKGSF	182
QY	184	KTYEV-NRKQVVALINSSSGTGLPKGVOLTHENLVTRFESHARDPIYGNVSPGTAIIIV	242
Db	183	VPLDVKNRKQVVALINSSSGTGLPKGVTRFTHGCAVTRFESHARDPIYGNVSPGTAIIIV	242
QY	243	VPRFHGGMFTTIGYLCGRFIWLTJTFDEDETLTKTLDYKCGSSVILVPTLFAILNRSFL	302
Db	243	VPRFHGGMFTTIGYLCGRFIWLTJTFDEDETLTKTLDYKCGSSVILVPTLFAILNRSFL	302
QY	303	LDKVDLSNLVEIASGSGAPLSKEIGEAARFRNLPYVQOGYGLTETTSAILITFEGDDKPG	362
Db	303	IDKFDLSNLVEIASGSGAPLSKEIGEAARFRNLPYVQOGYGLTETTSAILITFEGDDKPG	362
QY	363	ASGVVPLFPAKVIDLDTTKTKTLGPNRGEVCYVQPMKRGYVNDPEATRELIDEGCLHT	422
Db	363	ASGVVPLFPAKVIDLDTTKTKTLGPNRGEVCYVQPMKRGYVNDPEATRELIDEGCLHT	422
QY	423	GDIYDEEKEHFFIYDRLKSLIKKGYQVPALESVILQHPNIFDAGVAVGVPDPIAGEL	482
Db	423	GDIYDEEKEHFFIYDRLKSLIKKGYQVPALESVILQHPNIFDAGVAVGVPDPIAGEL	482
QY	483	PGAVVVLKKGKSMTEKEVMYDVASOVSNARKLRGCVRFVDEYKGLTGKIDKRAIREILK	542
Db	483	PGAVVVMKKGKSMTEKEVIVDVNSQVYVNHKRLRGVRFVDEYKGLTGKIDAKVIREILK	542
QY	543	KPVAKM 548	
Db	543	KPVAKM 548	
RESULT	4		
ID	Q26304	PRELIMINARY;	PRT; 548 AA.
AC	Q26304;		
DT	01-NOV-1996 (TRMBLrel. 01, Created)		
DT	01-NOV-1996 (TRMBLrel. 01, Last sequence update)		
DE	01-JUN-2001 (TRMBLrel. 17, Last annotation update)		
DE	LDCIFERASE.		
OS	Luticola mningrelica (Southern Russian firefly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;		
OC	Elateriformia; Cantharoidea; Lampyridae; Luciola.		
OX	NCBI_taxid=27446;		
RN	11]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=93277944; PubMed=8504162;		
RA	Devine J.H., Kutuzova G.D., Green V.A., Ugarova N.N., Baldwin T.O.;		
RT	"Luciferase from the east European firefly <i>Luciola mningrelica</i> : cloning		
RT	and nucleotide sequence of the cDNA, overexpression in <i>Escherichia</i>		
RL	<i>coli</i> and purification of the enzyme.";		
RL	Biochim. Biophys. Acta 1133:121-132(1993).		
DR	EMBL; S61961; AAB26932.1; -		
DR	HSSP; P08659; 1BA3.		
DR	InterPro: IPR000873; AMP-bind.		
DR	Pfam; PF00501; AMP-binding.1		
DR	PROSITE; PS00455; AMP-BINDING; UNKNOWN_1.		
SEQ	SEQUENCE 548 AA: 60494 MW: 62C8953BDFEBC423 CRC64;		

Query Match	82.7%;	Score 2335.5;	DB 5;	Length 548;
Best Local Similarity	81.7%;	Pred. No. 2.5e-153;		
Matches 446;	Conservative 47;	Mismatches 52;	Indels 1;	Gaps 1

Oy	4	MEDENLVYBPPEFFYPTEEGSSAGADLRKXMDRYAKLSALATNALGVYVYAEVLENSC	63
Oy	3	MEKENVYGPDPFPYPIEEGSSAGIQCHKXWYQAKLGAIAFSAALGVDSISOEYFDITC	62
Oy	64	CLGALKNGLVYDGRITACISENCEEFFIPVLAGLTIGVGAAPTNETIYLRLREYLSIGS	123
Db	63	RLAFAAMFPMKREBENHIALCSENCEEFFIPVLAGLTIGVGAAPTNETIYLRLREYLSIGTA	122
Oy	124	KPTTFSSKKGDLKVITYQKTAIVTAIKTIVILDSKVYRGYOSMDFIKNTPQFGKSSF	183
Oy	123	OPTIVFSRRRGPKVLEEVQKTVTCIKKIYVILDSKVNGGCHDEMEPIKKHVELFQPSF	182
Oy	184	KTVFV-NRKEQVALIMSSSTGLPKGVOLTHENLITFRSHADPTYGNOVSPGTAILLV	242
Db	183	VPIDVKNRKHVALIMNSSGSTGIPKGVRLTHGCAVTRSHADPTYGNOVSPGTAILLV	242
Oy	243	VPFHNGFMETTLGAYLLCGGRIVMLTKFDEEFLKTLQDYKCSSVILVPTFAILNRSFL	302
Db	243	VPFHNGFMETTLGAYFACQGRVYMLTKFDEEFLRLTLQDYKCTSVILVPTFAILNKSFL	302
Oy	303	LDKTDLSNLVFIASGAPRLSKELGEAVARFNP.PGVQKSGLTETTSALITTPREGDKG	362
Db	303	IDKFDLSNLVFIASGAPRLKVEGEAVARFNP.PGVQKSGLTETTSAPLITTPREGDKFG	362
Oy	363	ASGVVPLFKAVKVIDLDTKTKTLGPNRGEQCVGPMIMKGYVONPAPATREIIDEEGMLHT	422
Db	363	ASGVVPLFKAVKVIDLDTKTLGPNRGEICVKGPSLMLGYSNNPATRETIIDEEGMLHT	422
Oy	423	GDIGYDEEKHFVFDYRLKSLITKKGQVPALESLVLOHPNIPAGVAGVDPDLAGSL	482
Db	423	GDIGYDEDEHFFVYDRKSLITKKGQVPALESVLLOHPNIPAGVAGVDPDPADEL	482
Oy	483	PGAVVVLKKGSKMTEKEEMDVVASQVSNARKLRGVAFVDEVYKGLTGIDKAIAREILK	542
Db	483	PGAVVVMKMGKTMTEKEELVDVNSQVYNHNRKLRGVAFVDEVYKGLTGIDAKVIREILK	542
Oy	543	KPVAKM 548	
Db	543	KPOAKM 548	
RESULT 5			
ID	Q27688	PRELIMINARY;	PRT: 547 AA.
AC	Q27688;		
DT	01-NOV-1996 (TReMBLrel. 01, Created)		
DT	01-NOV-1996 (TReMBLrel. 01, Last sequence update)		
DT	01-JUN-2001 (TReMBLrel. 17, Last annotation update)		
DE	PHOTINUS-LUCIFERIN 4-MONOOXYGENASE (EC 1.13.12.7) (PHOTINUS-LUCIFERIN		
DE	4-MONOOXYGENASE (ATP-HYDROLYSING) (FIREFLY LUCIFERASE) (LUCIFERASE).		
OS	Lampyris noctiluca.		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;		
OC	Elateriformia; Cantharoidea; Lampyridae; Lampyris.		
OX	NCBI_TaxID=41311;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE-LIGHT ORGAN;		
RX	MEDLINE=96190714; PubMed=8611152;		
RA	Sala-newby G.B., Thomson C.M., Campbell A.K.;		
RT	"Sequence and biochemical similarities between the luciferases of the		
RT	glow-worm Lampyris noctiluca and the firefly Photinus pyralis.";		
RL	Biochem. J. 313:761-767(1996).		
CC	-I- CATALYTIC ACTIVITY: PHOTINUS LUCIFERIN + O(2) + ATP = OXIDIZED		
CC	PHOTINUS LUCIFERIN + CO(2) + H(2)O + AMP + DIPHOSPHATE + LIGHT.		
EMBL	X89479; CA61668.1; -.		
DR	HSSP; P08659; IBA3.		
DR	InterPro; IPR000873; AMP-bind.		
DR	InterPro; IPR000215; Serpin.		

DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 DR PROSITE: PS00284; SERPIN; UNKNOWN_1.
 KW Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 547 AA; 60365 MW; 8CB5653E0780D3EB CRC64;

Query Match 69.6%; Score 1965.5; DB 5; Length 547;
 Best Local Similarity 69.0%; Pred. No. 1e-127;
 Matches 372; Conservative 76; Mismatches 90; Indels 1; Gaps 1;

QY 4 MENDENTVGPPEPPYPIEEGSAGAQLRKYMDRAKL-GAIAFTNALTGVDYTAETLEKS 62
 DB 1 MEDAKNIMHGAPAPPYPIEDDTAGELKAKRKRYAQAQVPGTIAFTDAHEVNITAEYSEMA 60
 QY 63 CCGEALKNKGLVYDGAIALSCENCEEFTIPVLAGLFIGVAVPTNIYTLRELVSIGI 122
 DB 61 CRLAEPMKRRGLQNHIANVCENSLOFPMVCGALFIGVAVSTNDIYNERELYSLSI 120
 QY 123 SKPTIVSSKKGLDKYITVOKTVAIKTIVILDSKYDVRGYSMDNFIKKNTPOGFGSS 182
 DB 121 SPTIVSVSKRALOKILIGVOKKLPIDKIIILDSREDYMGKOSWSPFESHLPAGFNEYD 180
 QY 183 FKTEVNRKQVALINMSGSTGLPKGVOLTHENLYTRFSHARDPIYGNVSPGTALITV 242
 DB 181 YIPDSFDRKATALINMSGSTGLPKGVOLTHENLYTRFSHARDPIYGNVSPGTALITV 240
 QY 243 VPRHNGFMETTLGYLTCGRIVMLKTFDEETFLKTLQDYKSSVILVPLFALINSEL 302
 DB 241 IPRHNGFMETTLGYLTCGRIVMLKTFDEETFLKTLQDYKSSVILVPLFALINSEL 300
 QY 303 LDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGGLTETTSAILITPEGDDKPG 362
 DB 301 VDYDLSNLHEIASGAPLSKEIGEAVARFNLPGVROGGLTETTSAILITPEGDDKPG 360
 QY 363 ASGVVPLPKAKYIDDTKKTLCGRNREGEVCGPMLMGVYDNPETREIIDEEGMLHT 422
 DB 361 ASGVVPLPKAKYIDDTKKTLCGRNREGEVCGPMLMGVYDNPETREIIDEEGMLHT 420
 QY 423 GDIGYDEEHFEITVRLKSLIKYKGYQVPALESVLLQHPNIPFAGVAGVDDPITAGEL 482
 DB 421 GDIAAYDDEHFEITVRLKSLIKYKGYQVPALESVLLQHPNIPFAGVAGVDDPITAGEL 480
 QY 483 PGAVVVLKKGKSTKEKVDYVYASQVSNKRLRGVRFVDEVPKGLTGKIDGKAIREIL 541
 DB 481 PAAVVVLKKGKSTKEKVDYVYASQVSNKRLRGVRFVDEVPKGLTGKIDGKAIREIL 539

RESULT 6
 ID 027755 PRELIMINARY; PRT; 550 AA.

AC 027755;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE FIREFLY LUCIFERASE (EC 1.13.12.7) (PHOTINUS-LUCIFERIN 4-MONOOXYGENASE
 (ATP-HYDROLYSING)).
 GN LUC.
 OS Photinus pyralis (North American firefly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Elateriformia; Cantharoidea; Lampyridae; Photinus.
 OX NCBI_TaxID=7054;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AUTOGRAPHIA CALIFORNICA NUCLEAR POLYHEDROSIS VIRUS-SSE STRAIN
 RA Croizier G.;
 RL Submitted (FEB-1995) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: PHOTINUS LUCIFERIN + O(2) + ATP -> OXIDIZED
 CC PHOTINUS LUCIFERIN + CO(2) + H(2)O + AMP + DIPHOSPHATE + LIGHT.
 DR EMBL; X84846; CA59281.1; -;
 DR HSSP; P08659; 1BA3.

DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 KW Oxidoreductase.
 SQ SEQUENCE 550 AA; 60731 MW; 0A1CC749D0C96ADD CRC64;

Query Match 68.8%; Score 1941.5; DB 5; Length 550;
 Best Local Similarity 68.0%; Pred. No. 4.0e-126;
 Matches 368; Conservative 75; Mismatches 97; Indels 1; Gaps 1;

QY 4 MENDENTVGPPEPPYPIEEGSAGAQLRKYMDRAKL-GAIAFTNALTGVDYTAETLEKS 62
 DB 1 MEDAKNIMHGAPAPPYPIEDDTAGELKAKRKRYAQAQVPGTIAFTDAHEVNITAEYSEMA 60
 QY 63 CCGEALKNKGLVYDGAIALSCENCEEFTIPVLAGLFIGVAVPTNIYTLRELVSIGI 122
 DB 61 VRLAEAKRKRYGLNHNIRIVCENSLOFPMVCGALFIGVAVPANDIYNERELYSMSI 120
 QY 123 SKPTIVSSKKGLDKYITVOKTVAIKTIVILDSKYDVRGYSMDNFIKKNTPOGFGSS 182
 DB 121 SPTIVSVSKRALOKILIGVOKKLPIDKIIILDSKIDYQSFQSYTFVISHLPAGENEYD 180
 QY 183 FKTEVNRKQVALINMSGSTGLPKGVOLTHENLYTRFSHARDPIYGNVSPGTALITV 242
 DB 181 FVPESEFDRKATALINMSGSTGLPKGVOLTHENLYTRFSHARDPIYGNVSPGTALITV 240
 QY 243 VPRHNGFMETTLGYLTCGRIVMLKTFDEETFLKTLQDYKSSVILVPLFALINSEL 302
 DB 241 VPRHNGFMETTLGYLTCGRIVMLKTFDEETFLKTLQDYKSSVILVPLFALINSEL 300
 QY 303 LDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGGLTETTSAILITPEGDDKPG 362
 DB 301 IDKYDLSNLHEIASGAPLSKEIGEAVARFNLPGVROGGLTETTSAILITPEGDDKPG 360
 QY 363 ASGVVPLPKAKYIDDTKKTLCGRNREGEVCGPMLMGVYDNPETREIIDEEGMLHT 422
 DB 361 ASGVVPLPKAKYIDDTKKTLCGRNREGEVCGPMLMGVYDNPETREIIDEEGMLHT 420
 QY 423 GDIGYDEEHFEITVRLKSLIKYKGYQVPALESVLLQHPNIPFAGVAGVDDPITAGEL 482
 DB 421 GDIAAYDDEHFEITVRLKSLIKYKGYQVPALESVLLQHPNIPFAGVAGVDDPITAGEL 480
 QY 483 PGAVVVLKKGKSTKEKVDYVYASQVSNKRLRGVRFVDEVPKGLTGKIDGKAIREIL 542
 DB 481 PAAVVVLKKGKSTKEKVDYVYASQVSNKRLRGVRFVDEVPKGLTGKIDGKAIREIL 540
 QY 543 K 543
 DB 541 K 541

RESULT 7
 ID 027758 PRELIMINARY; PRT; 550 AA.

AC 027758;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE FIREFLY LUCIFERASE (EC 1.13.12.7) (PHOTINUS-LUCIFERIN 4-MONOOXYGENASE
 (ATP-HYDROLYSING)).
 GN LUC.
 OS Photinus pyralis (North American firefly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Elateriformia; Cantharoidea; Lampyridae; Photinus.
 OX NCBI_TaxID=7054;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AUTOGRAPHIA CALIFORNICA NUCLEAR POLYHEDROSIS VIRUS-SSE STRAIN
 RA Croizier G.;
 RL Submitted (FEB-1995) to the EMBL/Genbank/DBJ databases.


```
DR EMBL/D25415: BAA05005.1; -.
DR HSSP: P08659; 1BA3.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00455; AMP BINDING; 1.
SQ SEQUENCE 552 AA; 61000 MW; 85C14ED52B55366A CRC64
```

Query Match	64.6%;	Score 1825;	DB 5;	Length 552;
Best Local Similarity	62.5%;	Pred. No. 5.3e-118;		
Matches 338;	Conservative 92;	Mismatches 109;	Indels 2;	Gaps 2

[illegible]

RESULT	12			
027757				
ID	Q27757	PRELIMINARY;	PRT;	545 AA.
AC	Q27757;	002653;		
DT	01-NOV-1996	(TREMBLrel. 01, Created)		
DT	01-JUN-1997	(TREMBLrel. 04, Last sequence update)		
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)		
DE	LUCIFERIN 4-MONOMOXYGENASE (EC 1.13.12.7) (LUCIFERASE).			
OS	Photuris pennsylvanica.			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;			
OC	Elateriforma; Cantharoidea; Lampyridae; Photuris.			
OX	NCBI_TaxID=41716;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=LANTERN;			
RX	MEDLINE=97307756; PubMed=9165098;			
RA	Ye L., Buck L.M., Schaeffer H.J., Leach F.R.;			
RT	"Cloning and sequencing of a cDNA for firefly luciferase from Photuris			
RT	pennsylvanica."			

RT	Biochim.Biophys. Acta 1339:39-52(1997).
RL	-1- FUNCTION: PRODUCES GREEN LIGHT WITH A WAVELENGTH OF 562 NM (BY SIMILARITY).
CC	-1- CATALYTIC ACTIVITY: PROTHINUS LUCIFERIN + O(2) + AMP = OXIDIZED PROTHINUS LUCIFERIN + CO(2) + H(2)O + AMP + DIPHOSPHATE + LIGHT.
CC	-1- COFCTOR: REQUIRES A MAGNESIUM ION (BY SIMILARITY).
CC	-1- SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
CC	-1- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT COVALENT BINDING OF AMP TO THEIR SUBSTRATE.
DR	EMBL: U31240; AAB60897.1; -.
DR	HSSP: P08659; ILCI.
DR	InterPro: IPR000873; AMP-bind.
DR	Pfam: PF00501; AMP-binding; 1.
DR	PROSITE: PS00455; AMP-BINDING; UNKNOWN.1.
KM	Oxidoreductase; Monooxygenase; Photoprotein; Luminescence; Magnesium Peroxisome.
FT	SITE 543 545 MICROBODY TARGETING SIGNAL (POTENTIAL).
FT	SITE 545 545 MICROBODY TARGETING SIGNAL (POTENTIAL).
SD	SEQUENCE 545 AA: 60649 MW: 105682.8047CD6E CR664.

Query Match	57.5%;	Score 1623.5;	DB 5;	Length 545;
Best Local Similarity	56.5%;	Pred. No. 4.6e-104;		
Matches 307;	Conservative 97;	Mismatches 138;	Indels 1;	Gaps 1.

QY	7	DEINIYGEPEFPIEDIGSGAOLRKYMDRYAKUL-GAIAFNALITGVDIYAYATKESCOL	65
Db	3	DKNITLGPREFPIHLDGTADGBDMFALSRVADISGCALTALNATKRENVLYEERFALSCRL	62
QY	66	GEALKNYGLVDBGRIALCSENCPEEFPIVLAGLEFIGVAPNTEIYTLRELVHSLGISKP	12
Db	63	AESFRRKQKJMDTTIAVNCSENGLOFEPLIASLITGLIAPASDKYERELLHSLGIYKP	12
QY	126	TIYFSSKGLDKVIYQKTVNLAIKTIVILDSKVDRBGYSMDNFIKKPNPOGFKSSSKRT	18
Db	123	RIIFCSKNTFQOKLVNKSXKLYVETIIILDLNEDLGGYQCLNFIQNSDINLDVKRKP	18
QY	186	VEVNRKEOVALINNSSGSGTGLPGVOLFNHLYTRFSHADPTIYGNOVSPGATILTVPF	24
Db	183	NSFNREDOVALWFESSGTTGVSKGMYLTKHKIYARFESHCKDPTFGNAINPTAILTLVIF	24
QY	246	HNGFGMTTLGYLTGCFRIYMLTKPDEETPLFKTLQDPCSSVILVPTLPAIINRSELDK	30
Db	243	HNGGMTTLTGFTGCFRVALMHPEERKLFQSLQDQKVVESTILLVPTLMAFPKSAVLEK	30
QY	306	YDLSNLEVLTASGAPLSKTEIGEAIVARNLPDVGQYGLTETTSALITTPREGDDKPGASG	36
Db	303	YDLSHLEKLTASGAPLSKEIGEMVKKRKLNFVROGSLGETTSAVLITPDDVDVRGSGTG	36
QY	366	KVYPLFRAXYIDLDTKKTGLRNRGEVONVGRPMILMKRYUNPEATREIILDEBGSJHTGDI	42
Db	363	KIYFNAHVAKVYDPTTGKLLIGRPETLETFKQDMITMKSYUNNEEATYTAIITINKOGMLRSGDI	42
QY	426	GUYDEEKNFEIYDRLKSLIKYKGVOVPRAELSEVLLQHPNIFDAGVAGVDDPIAGELPQA	48
Db	423	AYYDNGSHFEIYDRLKSLIKYKGVOYARALIEGLILQHPITYVAGVTGIPDEAAGELPRA	48
QY	486	VVYLKKGSKTEKEVMDVYASOVNSAKFLRGVRFVDEVPKGLTGKIDGAIAREILAKRPV	54
Db	483	GVVYQTKETYLEQNIQVNEFSSOVSTAKMLRGVGFDEFDEIFKSGTGKIDRKRVLROMEKHK	54
QY	546	AKM 548	
Db	543	SKL 545	

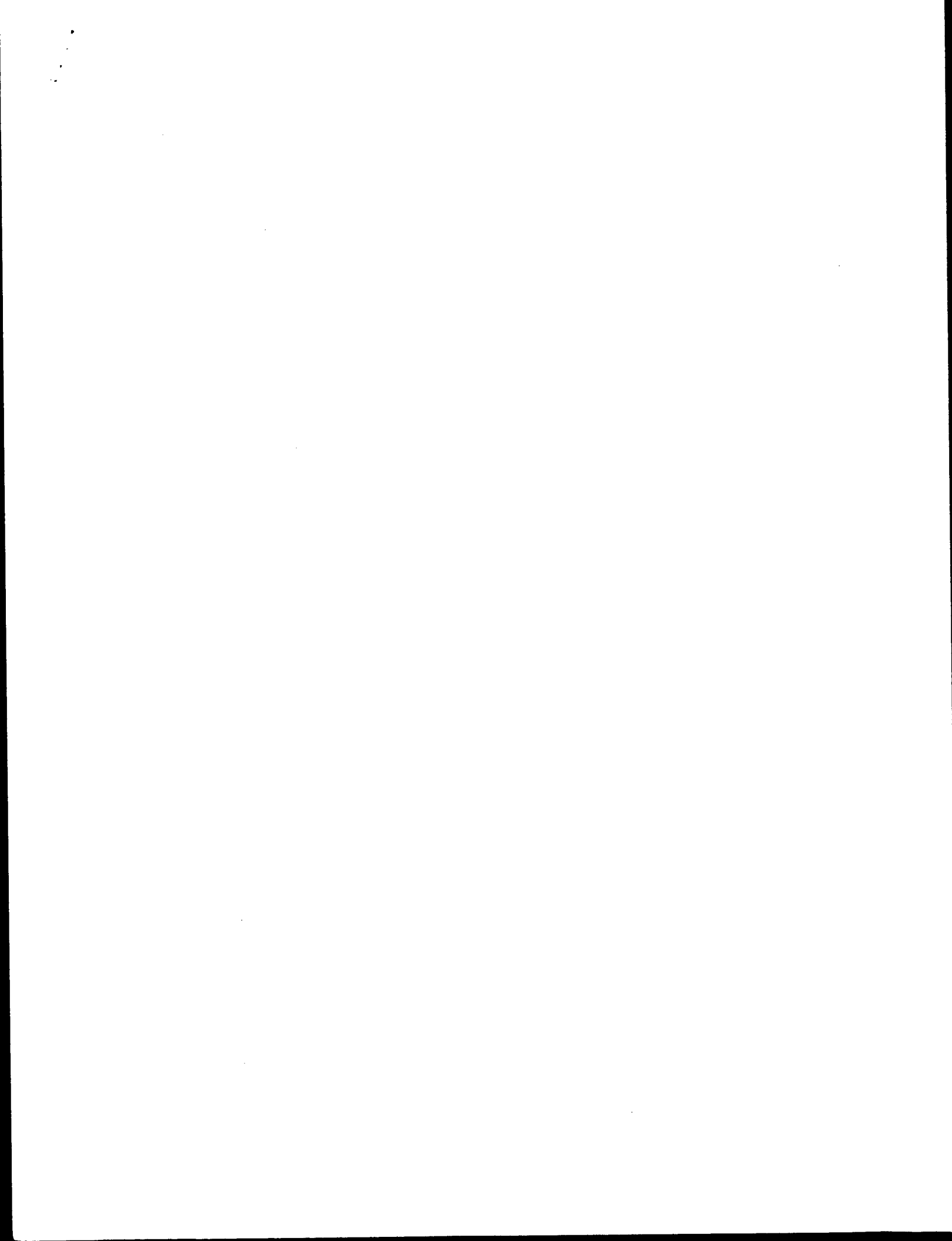
RESULT	13	
Q9U4U8		
ID	Q9U4U8	PRELIMINARY;
AC	Q9U4U8;	PRT; 545 AA.
DT	01-MAY-2000	(TREMBLrel. 13, Created)
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)
DE	LOCIFERASE.	


```

Db      124 KPRIFASITITDRVAKVAAKMKFYKGITIALSGTSKRRKNITDYKLEMEDEKFTQP--- 186
QY      179 KSSPKTEVANKKEQVALIMNMSGSTGLPKGYOLTHEMLYTRSHARDPIYGNQVSP-- 235
Db      181 ---DPTSPANKDEBVSILVSCSGTGTLPKPVOLTMNLATLD-----SOIQPTVI 229
QY      236 ---GTAILTVVPEHHGFMFTTLGYLTGCRFIMVLTKRDEETFFLTQDYKCSSVILPT 292
Db      230 PMEEVTLITLVIPIWFIHAFGLTLITACVAGALVYLLPKREKFLISALIEKRVMMAFVPP 269
QY      293 LEAIIINRELLDKYDLSNLVELASGAPLSKEISEIVAAARRNLPFGVNOGGLTETSATIT 352
Db      290 LMFVLAHPRIYDKDLSLMLILCGAPLSKEDTDQIKERIGVPIKOGYGLSESTLSVL 349
QY      333 ITPEGDDKPGASGVAVPLFKAVIDLDTKKTIGNRRGEVCKPMLMKGYVDNPEATRE 412
Db      350 VONNEFCRPGSGVGLKVGIVAKVIDPDTGKTLGANEGERGELCFKDDGIMKGYIDGTSTQF 409
QY      413 IIDEGMLHTDIDIGYDEBEKHEFFIVDRLSKLIKXGYGVPAELSESLDHPNIFAGVA 472
Db      410 AI-KDGMHTSDIDIGYDDDEFFIVDRIKELIKKGIQVPAELFALLTNDXIKIDAAV 468
QY      473 GVPPPIGELPGAVYVLKKGKSMTEKEVMDIVASQVSNARKLRGVRVDEVPKGLTGKI 532
Db      469 GRPPEAGEELPLAVVQAQVOLTENEYIOFVNDNAPSARKLRGVIIVDEIPKNSGKI 528
QY      533 DGKAIRIELIKRPVAKM 548
Db      529 LRRILREMLKKQSKL 544

```

```
Search completed: November 29, 2001, 01:59:32
Job time: 377 sec
```



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2001, 01:59:32 ; Search time 64.96 Seconds
(without alignments)
1233.948 Million cell updates/sec

Title: US-09-581-241-6

Perfect score: 2823
Sequence: I MNNNDENIYVGPPEPPYPL.....TGKIDGKAIRKILKKPVAKM 548

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: SP.TREMBL.17:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2793	98.9	548	5	Q27321 luciola lat
2	2792	98.9	548	5	Q27348 luciola lat
3	2343.5	83.0	548	5	Q25118 notaria par
4	2335.5	82.7	548	5	Q26304 luciola min
5	1967.5	69.7	547	5	Q27688 lampyrus no
6	1941.5	68.8	550	5	Q27755 photinus py
7	1918.5	68.0	550	5	Q27758 photinus py
8	1900	67.3	548	5	Q9GPP9 pyrocoelia
9	1892	67.0	548	5	Q26076 pyrocoelia
10	1831	64.9	552	5	Q94697 photuris pe
11	1825	64.6	552	5	Q94656 photuris pe
12	1625.5	57.6	545	5	Q27757 photuris pe
13	1457.5	51.6	545	5	Q94048 photuris pe
14	1332	47.2	546	5	Q94047 photuris pe
15	1023	36.2	544	5	Q9VC66 drosophila
16	775.5	27.5	546	10	Q9C5H2 aradidopsi
17	771	27.3	636	10	Q42879 lithosperu
18	770	27.3	562	10	Q9FGW4 aradidopsi
19	756	26.8	557	10	Q48869 populus dal

20	754	26.7	544	10	Q9M0X9 aradidopsi
21	753.5	26.7	548	10	Q48868 populus bal
22	748.5	26.5	580	10	Q9LMV8 aradidopsi
23	746	26.4	544	10	Q9LLS0 rubius idae
24	742	26.3	543	10	Q9LL49 rubius idae
25	742	26.3	585	10	Q42880 lithosperu
26	738.5	26.2	542	10	Q42943 nicotiana t
27	737.5	26.1	597	5	Q9Y158 drosophila
28	732.5	25.9	535	10	Q81139 populus tre
29	732.5	25.9	545	10	Q9SMW8 solanum tub
30	720	25.5	544	5	Q19339 caenorhabdi
31	719	25.5	1549	10	Q9LMW7 aradidopsi
32	714	25.3	522	2	Q9K3W1 streptomyce
33	713	25.3	570	10	Q81140 populus tre
34	691	24.5	556	10	Q9M7S2 iolium pere
35	682	24.2	570	10	Q9LU36 aradidopsi
36	674.5	23.9	542	10	Q9LQ12 aradidopsi
37	673.5	23.9	570	10	Q9M7S3 iolium pere
38	672	23.8	591	10	Q9LL51 rubius idae
39	664.5	23.5	542	10	Q9FOY7 capsicum an
40	655	23.2	557	10	Q9M7S1 iolium pere
41	647	22.9	544	10	Q9FE44 aradidopsi
42	635	22.5	552	1	Q28762 archaeoglob
43	632.5	22.4	566	10	Q49414 aradidopsi
44	600	21.3	566	2	Q9K8A2 bacillus ha
45	597.5	21.2	569	1	Q28502 archaeoglob

ALIGNMENTS

RESULT 1
ID Q27321 PRELIMINARY: PRT: 548 AA.
AC Q27321;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DI 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE LUCIPERASE.
OS Luciola lateralis (firefly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Elateriforma; Cantharoidea; Lampyridae; Luciola.
OX NCBI_TaxID=7052;
RN [1]
RP SEQUENCE FROM N.A.
RA Cho K., Choi Y., Boo K.;
RL Submitted (FEB-1996). to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Cho K.H., Choi Y., Boo K.;
RL Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN-MUJU; TISSUE-ABDOMEN;
RC Cho K., Choi Y., Boo K.;
RL Submitted (MAY-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL: Z69619; CA93444.1; -;
DR EMBL: U49181; AA91471.1; -;
DR EMBL: U51019; AAB00229.1; -;
DR HSSP: P08659; ILCT.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00455; AMP-BINDING; 1.
SQ SEQUENCE 548 AA: 60048 MW: E26CEC26FA23E5E8 CRC64;

Query Match 98.9%; Score 2793; DB 5; Length 548;
Best Local Similarity 98.7%; Pred. No. 1.2e-184;
Matches 541; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
OY I MNNNDENIYVGPPEPPYPLIEGSAQALRKYMDRYAKLGAIFATNALGCVDTTAYEIE 60

Db	1	MEMDNDENITVYGPAPFVPEDEEGSAGADLRKMRMYALGAIATFNALTGDVITYAEYLE	60
QY	61	KSCCGLEAKNTGLVYVDGRIALCSNCEEFIPVLAGLFIVGVAAPIINEIYTLRELVHSL	120
Db	61	KSCCGLEAKNTGLVYVDGRIALCSNCEEFIPVLAGLFIVGVAAPIINEIYTLRELVHSL	120
QY	121	GISKEPIYSSKKGLDKYITVQKTTATLKTIVILIIDSKYDVRGYSMDNFIKKNPQGFKG	180
Db	121	GISKEPIYSSKKGLDKYITVQKTTATLKTIVILIIDSKYDVRGYSMDNFIKKNPQGFKG	180
QY	181	SSFKEIVENRKAQVALINSSSGSTGLPKGVOLTHENIVTRFSHARDEPIYGNVSPGTAIL	240
Db	181	SSFKEIVENRKAQVALINSSSGSTGLPKGVOLTHENIVTRFSHARDEPIYGNVSPGTAIL	240
QY	241	TVVPPHHGGMFTTIGYLTGCFRIYMLTKPEEFILTIODYKCSSVLVPTLFTALINRS	300
Db	241	TVVPPHHGGMFTTIGYLTGCFRIYMLTKPEEFILTIODYKCSSVLVPTLFTALINRS	300
QY	301	ELLDKYDLSNIVEIASGAPLSKEITGEAVARFNLPGVROGYGLTETTSNAIITPEGDCK	360
Db	301	ELLDKYDLSNIVEIASGAPLSKEITGEAVARFNLPGVROGYGLTETTSNAIITPEGDCK	360
QY	361	PGASKQVPLPKAKYIDIDTKKTLGPNRNGVCYKGMINKGYDNEARREIIDEEGWL	420
Db	361	PGASKQVPLPKAKYIDIDTKKTLGPNRNGVCYKGMINKGYDNEARREIIDEEGWL	420
QY	421	HTGDIGYDDEKHFEFIVDRLSLIRKYGQYVPAELESVLLQHPNIFPDAGVAVPPPIAG	480
Db	421	HTGDIGYDDEKHFEFIVDRLSLIRKYGQYVPAELESVLLQHPNIFPDAGVAVPPPIAG	480
QY	481	ELPGVAVVLLKKGSKMTEKEVNDVYASOVSNAKRLRGVRRVDEVPKGLTGKIDGKAIRREI	540
Db	481	ELPGVAVVLLKKGSKMTEKEVNDVYASOVSNAKRLRGVRRVDEVPKGLTGKIDGKAIRREI	540
QY	541	LKKPVAKM 548	
Db	541	LKKPVAKM 548	
RESULT	2		
Q27348			
ID	Q27348	PRELIMINARY;	PRT; 548 AA.
AC	Q27348;		
DT	01-NOV-1996 (TREMblrel. 01, Created)		
DT	01-NOV-1996 (TREMblrel. 01, Last sequence update)		
DT	01-JUN-2001 (TREMblrel. 17, Last annotation update)		
DE	LUCIFERASE.		
OS	Luciola lateralis (Firefly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;		
OC	Elateriforma; Cantharidoidea; Lampyridae; Luciola.		
NCBI_Taxid=7052;			
LN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-MUTU; TISSUE=ABDOMEN;		
RC	Cho K.H., Choi Y., Boo K.;		
RA	Submitted (JUN-1995) to the EMBL/Genbank/DDBJ databases.		
RL	EMBL; U49182; AAA91472.1; -;		
DR	EMBL; Z49891; CAA90072.1; -;		
DR	HSSP; P08659; IBA3.		
DR	InterPro; IPR000873; AMP-bind.		
DR	Pfam; PF00501; AMP-binding; 1.		
DR	PROSITE; PS00455; AMP_BINDING; 1.		
QO	SEQUENCE 548 AA; 60062 MW; D237FE6C97CA3B8D CRC64;		

Query Match 98.9%; Score 2792; DB 5; Length 548,

Best Local Similarity 98.7%; Pred. No. 1,4e-184;
Matches 541; Conservative 3; Mismatches 4; Indels 0; Gaps 0

QY	61	KSCCGEALKNYGLVVDGRIALCSCNCEEFPLVAGLFGVGAAPNNEYTLRELHSL	120
Db	61	KSCCGEALKNYGLVVDGRIALCSCNCEEFPLVAGLFGVGAAPNNEYTLRELHSL	120
QY	121	GISKEPTIVFSSKGLDKVIYVQKTYTAIKTVIILDSKVDYKGYSDMNFTKNTPGFGK	180
Db	121	GISKEPTIVFSSKGLDKVIYVQKTYTAIKTVIILDSKVDYKGYSDMNFTKNTPGFGK	180
QY	181	SSFKEIVENRKROVALIMNSSGSTGLPKGYOLTHEMNTFRSHARDPIYGNVSPGAIL	240
Db	181	SSFKEIVENRKROVALIMNSSGSTGLPKGYOLTHEMNTFRSHARDPIYGNVSPGAIL	240
QY	241	TVVPFHGFGMTTLGYLTGFRIVMLTKPFDEEFLKTLDPYKCSSYLVPFLAILNRS	300
Db	241	TVVPFHGFGMTTLGYLTGFRIVMLTKPFDEEFLKTLDPYKCSSYLVPFLAILNRS	300
QY	301	ELLDYKDSLNLVEIASGAPLPSKEIGEVANRFMLPGVROGYGLTETTSAILITPBGDK	360
Db	301	ELLDYKDSLNLVEIASGAPLPSKEIGEVANRFMLPGVROGYGLTETTSAILITPBGDK	360
QY	361	PGASGRVPLERAKVIDIDTKTKTGLPNRNGEVCYKGMILKKGYDNPENAREITIDEGM	420
Db	361	PGASGRVPLERAKVIDIDTKTKTGLPNRNGEVCYKGMILKKGYDNPENAREITIDEGM	420
QY	421	HGDIIGYDEKHEFFIYDRLSLIKRYGYVPPAELESYLLQHPNITFDAGVAGVPPDAG	480
Db	421	HGDIIGYDEKHEFFIYDRLSLIKRYGYVPPAELESYLLQHPNITFDAGVAGVPPDAG	480
QY	481	ELPGAVVLLKKGKSMTEKEVMDYVAASOVNASAKRLRGSEVRVDEVPGLNGKIDGARIREI	540
Db	481	ELPGAVVLLKKGKSMTEKEVMDYVAASOVNASAKRLRGSEVRVDEVPGLNGKIDGARIREI	540
QY	541	LKKPYAKM 548	
Db	541	LKKPYAKM 548	
RESULT	3		
Q25118			
ID	Q25118	PRELIMINARY;	PRT; 548 AA.
AC	Q25118		
DT	01-NOV-1996 (TREMBlrel. 01, Created)		
DT	01-NOV-1996 (TREMBlrel. 01, Last sequence update)		
DT	01-JUN-2001 (TREMBlrel. 17, Last annotation update)		
DE	LOCIFERRASE.		
OS	Holaria parvula.		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;		
OC	Elaeteriformia; Cantharoidea; Lampyridae; Holaria.		
OX	NCBI_TaxID=39323;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=96061635; PubMed=7480137;		
RA	Ohmura Y., Ohba N., Toh H., Tsuji F.I.;		
RT	"Cloning, expression and sequence analysis of cDNA for the luciferase		
RT	from the Japanese fireflies, <i>Pyrococella miyako</i> and <i>Holaria parvula</i> ."		
RL	Photochem. Photobiol. 62:309-313(1995).		
DR	EMBL; L39929; AAC37253.1; -.		
DR	HSSP; P08659; ILCI.		
DR	InterPro; IPR000873; AMP-bind.		
DR	Pfam; PF00501; AMP-binding; I.		
DR	PROSITE; PS00455; AMP_BINDING; UNKNOWN_1.		
Q0	SEQUENCE 548 AA; 60365 MW; 9694F3EB46CAB85 CRC64;		

Query Match 83.0%; Score 2343.5; DB 5; Length 548

DR Pfam; PF00501; AMP-binding; 1.
 DR PROSITE; PS00455; AMP_BINDING; 1.
 DR PROSITE; PS00284; SERPIN; UNKNOWN_1.
 KW Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 547 AA; 60365 MW; 8CB5653E0780D3EB CRC64;

Query Match 69.7%; Score 1967.5; DB 5; Length 547;
 Best Local Similarity 69.0%; Pred. No. 1.3e-127;
 Matches 372; Conservative 76; Mismatches 90; Indels 1; Gaps 1;

QY 4 MENDNIYVGPPEFYPIEGSAGADLRKYMRYAKL-GAIAFTNALTGVDYTAAYLEKS 62
 DB 1 MEDAKNIKKGPAPFPPLDGTAGBOLHKAMKRYALPGIATPDHAENVITYAYEFEMA 60
 QY 63 COLGALNNYGLVNDGRALCSNCEEFPIYLAGLFTGVGAPAPNETYTLRELVHSLGI 122
 DB 61 CRLEETMRKRYGLQHNIAVCSNSLOPFMPVCGALFTGVGASTINDYNERELYNLSLI 120
 QY 123 SKPTIVSSKKGLDKVITVOKTVTAIKTIVILDSKVDRGVQSMDFTKKTPGFGSS 182
 DB 121 SQPTIVSSKKGLDKVITVOKTVTAIKTIVILDSKVDRGVQSMDFTKKTPGFGSS 180
 QY 183 FKTYEVNKRKEQVALIMNSGSTGLPKGYOLTHENIVTRFSHARDPIYGNVSPGTAIITV 242
 DB 181 YIPSEFDRKDTALIMNSGSTGLPKGYOLTHENIVTRFSHARDPIYGNVSPGTAIITV 240
 QY 243 VPHHNGFMFTTGLVLCGFRIVMLTKFDEETFLKTLQDYKCSVILVPTFAIINSEL 302
 DB 241 IPFHNGFMFTTGLVLCGFRIVMLTKFDEETFLKTLQDYKCSVILVPTFAIINSEL 300
 QY 303 LDKYDLSNLVLIASGAPLSEKEIGEAVARRNLPVGRVGYGLTETTSALITTPGDDKPG 362
 DB 301 IDKDYDLSNLVLIASGAPLSEKEIGEAVARRNLPVGRVGYGLTETTSALITTPGDDKPG 360
 QY 363 ASGVVPLFFKAKVIDLDTKTLGPNRGEVCKGPMLEKGVNDPEATREIIDEGLWLT 422
 DB 361 ACGRVVPFFSAKVIDLDTKTLGPNRGEVCKGPMLEKGVNDPEATREIIDEGLWLT 420
 QY 423 GDIGYDEEKKHFFIVDRLSLTKKGYOVPAELLESYLQHPNIFDAGVAGVPPIAGEL 482
 DB 421 GDIAVWDEDEHFFIVDRLSLTKKGYOVPAELLESYLQHPNIFDAGVAGVPPIAGEL 480
 QY 483 PGAVVVLKKGKSMTEKEMVDYVASOVNAKRLRGVRFVDEVPKGLGKIDGKAIREL 541
 DB 481 PAAVVVLEHGKTMTEKEIVDYVASOVNAKRLRGVRFVDEVPKGLGKIDGKAIREL 539

RESULT 6
 Q27755 PRELIMINARY; PRT; 550 AA.
 ID 027755;
 AC 027755;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE FIREFLY LUCIFERASE (EC 1.13.12.7) (PHOTINUS-LUCIFERIN 4-MONOOXYGENASE (ATP-HYDROLYSING)).
 GN LUC.
 OS Photinus pyralis (North American firefly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Plekoyota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Elateriformia; Cantharoidae; Lampyridae; Photinus.
 OX NCBI_Taxid=7054;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AUTOGRAHA CALIFORNICA NUCLEAR POLYHEDROSIS VIRUS-SSE STRAIN 1.2 RECOMBINANT 1.1.1;
 RA Croizier G.;
 Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: PHOTINUS LUCIFERIN + O(2) + ATP = OXIDIZED PHOTINUS LUCIFERIN + CO(2) + H(2)O + AMP + DIPHOSPHATE + LIGHT.
 DR EMBL; X84846; CAA59281.1; -;
 HSSE; P08659; IBA3.

DR InterPro; IPR000873; AMP-bind.
 DR Pfam; PF00501; AMP-binding; 1.
 DR PROSITE; PS00455; AMP_BINDING; 1.
 KW Oxidoreductase.
 SQ SEQUENCE 550 AA; 60731 MW; 0A1CC749D0C96ADD CRC64;

Query Match 68.8%; Score 1941.5; DB 5; Length 550;
 Best Local Similarity 68.0%; Pred. No. 8e-126;
 Matches 368; Conservative 75; Mismatches 97; Indels 1; Gaps 1;

QY 4 MENDNIYVGPPEFYPIEGSAGADLRKYMRYAKL-GAIAFTNALTGVDYTAAYLEKS 62
 DB 1 MEDAKNIKKGPAPFPPLDGTAGBOLHKAMKRYALPGIATPDHAENVITYAYEFEMA 60
 QY 63 COLGALNNYGLVNDGRALCSNCEEFPIYLAGLFTGVGAPAPNETYTLRELVHSLGI 122
 DB 61 VRLAEMKRYGLNTHRIIVCSNSLOPFMPVGLGALFTGVAAPANDIYNERELNSMNI 120
 QY 123 SKPTIVSSKKGLDKVITVOKTVTAIKTIVILDSKVDRGVQSMDFTKKTPGFGSS 182
 DB 121 SQPTIVSSKKGLDKVITVOKTVTAIKTIVILDSKVDRGVQSMDFTKKTPGFGSS 180
 QY 183 FKTYEVNKRKEQVALIMNSGSTGLPKGYOLTHENIVTRFSHARDPIYGNVSPGTAIITV 242
 DB 181 FVPESEFDRKDTALIMNSGSTGLPKGYOLTHENIVTRFSHARDPIYGNVSPGTAIITV 240
 QY 243 VPHHNGFMFTTGLVLCGFRIVMLTKFDEETFLKTLQDYKCSVILVPTFAIINSEL 302
 DB 241 VPHHNGFMFTTGLVLCGFRIVMLTKFDEETFLKTLQDYKCSVILVPTFAIINSEL 300
 QY 303 LDKYDLSNLVLIASGAPLSEKEIGEAVARRNLPVGRVGYGLTETTSALITTPGDDKPG 362
 DB 301 IDKDYDLSNLVLIASGAPLSEKEIGEAVARRNLPVGRVGYGLTETTSALITTPGDDKPG 360
 QY 363 ASGVVPLFFKAKVIDLDTKTLGPNRGEVCKGPMLEKGVNDPEATREIIDEGLWLT 422
 DB 361 GVGAVVPFFSAKVIDLDTKTLGPNRGEVCKGPMLEKGVNDPEATREIIDEGLWLT 420
 QY 423 GDIGYDEEKKHFFIVDRLSLTKKGYOVPAELLESYLQHPNIFDAGVAGVPPIAGEL 482
 DB 421 GDIAVWDEDEHFFIVDRLSLTKKGYOVPAELLESYLQHPNIFDAGVAGVPPIAGEL 480
 QY 483 PGAVVVLKKGKSMTEKEMVDYVASOVNAKRLRGVRFVDEVPKGLGKIDGKAIREL 542
 DB 481 PAAVVVLEHGKTMTEKEIVDYVASOVNAKRLRGVRFVDEVPKGLGKIDGKAIREL 540
 QY 543 K 543
 DB 541 K 541

RESULT 7
 Q27758 PRELIMINARY; PRT; 550 AA.
 ID 027758;
 AC 027758;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE FIREFLY LUCIFERASE (EC 1.13.12.7) (PHOTINUS-LUCIFERIN 4-MONOOXYGENASE (ATP-HYDROLYSING)).
 GN LUC.
 OS Photinus pyralis (North American firefly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Plekoyota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Elateriformia; Cantharoidae; Lampyridae; Photinus.
 OX NCBI_Taxid=7054;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AUTOGRAHA CALIFORNICA NUCLEAR POLYHEDROSIS VIRUS-SSE STRAIN 1.2 RECOMBINANT 8.1.1;
 RA Croizier G.;
 Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: PHOTINUS LUCIFERIN + O(2) + ATP = OXIDIZED
 CC PHOTINUS LUCIFERIN + CO(2) + H(2)O + AMP + DIPHOSPHATE + LIGHT.
 CC EMBL: X84847; CAA59282.1; -
 DR HSSP: P08659; 1BA3.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 DR Oxidoreductase.
 KM
 SQ SEQUENCE 550 AA; 60602 MW; D963B300D030F119 CRC64;

Query Match 68.0%; Score 1918.5; DB 5; Length 550;
 Best Local Similarity 67.5%; Pred. No. 3,1e-124;
 Matches 365; Conservative 76; Mismatches 99; Indels 1; Gaps 1;

QY 4 MENDENTVGPPEPPYPIEBSAGQRLKRYDRAKL-GAIAFNALGVDTYAEYLEKS 62
 DB 1 MEDAKNIKGPAPFYPLEDGTAGEOLHKAKRKRYALVPTIAFDHAEVITYAEYFEMS 60
 QY 63 CCLGEALKNYGLVVDGRITALCSENCEEFPIPVLAGLFIGVAVPNTIEYTLRELVSIG 122
 DB 61 VRLAEAKRKRIGLTNHRIIVCSNSLOFPMVPGALFIGVAVAPANDIYNERELLSMN 120
 QY 123 SKPTIVSSKKGLDKVITVQKTYATKTYIYILDSKVYRGYQSDNFIKNTPOGFKGSS 182
 DB 121 SQPTIVSVSKGLQKILNVQKLPITQKIIIMDSKTDYQGFQSMYTFVTSHPGFNEYD 180
 QY 183 FKTYEVNRKEQVALIMNSSGSTGLPKGVOLTHENITYRSHADPIYGNVSPCTALTY 242
 DB 181 FVPEFPRDRTIALIMNSSGSTGPKGVALPHRTACVRFSHADPIFGNDIIPDTALISV 240
 QY 243 VPFHNGFMFTTGLYLCGFRIVMLTKFDEFTLKTLDYDKCSSVILVPTLFAILNRSEL 302
 DB 241 VPFHNGFMFTTGLYLCGFRIVMLTKFDEFTLKTLDYDKCSSVILVPTLFAILNRSEL 300
 QY 303 LDKYDLSNVLASGAPLSKEIGEAVALRRPNLPVGRQYGTLETTSAILITPBGDDKP 362
 DB 301 IDKYDLSNVLASGAPLSKEIGEAVALRRPNLPVGRQYGTLETTSAILITPBGDDKP 360
 QY 363 ASGVVPLFAKVIDLDTKKTLPGRNRCGVCGPMLKGYVNDPEATREIIDEGLHT 422
 DB 361 AVGVVPLFAKVIDLDTKKTLPGRNRCGVCGPMLKGYVNDPEATREIIDEGLHT 420
 QY 423 GDIGYDEKHEFFIVDRLSLTKYGYVPAELESVLLQHPNIFDAGVAPPPPAGE 482
 DB 421 GDIGYDEKHEFFIVDRLSLTKYGYVPAELESVLLQHPNIFDAGVAPPPPAGE 480
 QY 483 PGAVVVLKKGKSMTEKEMDVYASQVSNARKLRGVRVDEVPKGLTGKIDGKAIREIL 542
 DB 481 PAAVVVLEHGKTMTEKEITVDYASQVTTAKKLRGVVFEVDEVPKGLTGKIDGKAIREIL 540
 QY 543 K 543
 DB 541 K 541

RESULT 8
 Q9GPF9 PRELIMINARY; PRT; 548 AA.

AC Q9GPF9; 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE LUCIFERASE.
 OS Pyrococcus rufa.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Elateriformia; Cantharoidea; Lampyridae; Pyrocoelia.
 NCBI_TaxID=71223;
 RX NCB1
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIGHT ORGAN;
 RA Lee K.S., Park H.J., Bae J.S., Lee K.S., Shon H.D., Jin B.R.;

RT "Molecular Cloning and Expression of cDNA encoding luciferase from the
 RT firefly, *Pyrococcus rufa*,"
 RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RL
 DR EMBL: AF328553; AAC645439.1; -
 DR InterPro: IPR000873; AMP-bind.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00284; SERPIN; UNKNOW.
 DR
 SQ SEQUENCE 548 AA; 60787 MW; FE86DB7DDDEFA9B CRC64;

Query Match 67.3%; Score 1900; DB 5; Length 548;
 Best Local Similarity 68.0%; Pred. No. 5.8e-123;
 Matches 367; Conservative 72; Mismatches 99; Indels 2; Gaps 2;

QY 4 MEND-ENTYGPPEPPYPIEBSAGQRLKRYDRAKL-GAIAFNALGVDTYAEYLEK 61
 DB 1 MEDDSKHHMHRHSILMEDGTAGEOLHKAKRKRYALVPTIAFDHAEVITYAEYFEMS 60
 QY 62 SCCLGEALKNYGLVVDGRITALCSENCEEFPIPVLAGLFIGVAVPNTIEYTLRELVSIG 121
 DB 61 SCRLAETMKRYGLGQHHIIVCSNSLOFPMVPGALFIGVAVAPANDIYNERELYSLF 120
 QY 122 ISKPTIVSSKKGLDKVITVQKTYATKTYIYILDSKVYRGYQSDNFIKNTPOGFKGSS 181
 DB 121 ISQPTIVSVSKGLQKILNVQKLPITQKIIIMDSKTDYQGFQSMYTFVTSHPGFNEY 180
 QY 182 FKTYEVNRKEQVALIMNSSGSTGLPKGVOLTHENITYRSHADPIYGNVSPCTALTY 241
 DB 181 DYIPDSFDEETALIMNSSGSTGLPKGVOLTHENITYRSHADPIYGNVSPCTALTY 240
 QY 242 VPFHNGFMFTTGLYLCGFRIVMLTKFDEFTLKTLDYDKCSSVILVPTLFAILNRSEL 301
 DB 241 VPFHNGFMFTTGLYLCGFRIVMLTKFDEFTLKTLDYDKCSSVILVPTLFAILNRSEL 300
 QY 302 LDKYDLSNVLASGAPLSKEIGEAVALRRPNLPVGRQYGTLETTSAILITPBGDDKP 361
 DB 301 LDKYDLSNVLASGAPLSKEIGEAVALRRPNLPVGRQYGTLETTSAILITPBGDDKP 360
 QY 362 ASGVVPLFAKVIDLDTKKTLPGRNRCGVCGPMLKGYVNDPEATREIIDEGLHT 421
 DB 361 ASGVVPLFAKVIDLDTKKTLPGRNRCGVCGPMLKGYVNDPEATREIIDEGLHT 420
 QY 422 GDIGYDEKHEFFIVDRLSLTKYGYVPAELESVLLQHPNIFDAGVAPPPPAGE 481
 DB 421 GDIGYDEKHEFFIVDRLSLTKYGYVPAELESVLLQHPNIFDAGVAPPPPAGE 480
 QY 482 LPGAIVVVLKKGKSMTEKEMDVYASQVSNARKLRGVRVDEVPKGLTGKIDGKAIREIL 541
 DB 481 LPGAIVVVLKKGKSMTEKEMDVYASQVSNARKLRGVRVDEVPKGLTGKIDGKAIREIL 540

RESULT 9
 Q26076 PRELIMINARY; PRT; 548 AA.

AC Q26076; 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE LUCIFERASE.
 OS Pyrococcus miyako.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Elateriformia; Cantharoidea; Lampyridae; Pyrocoelia.
 NCBI_TaxID=39363;
 RX NCB1
 RP SEQUENCE FROM N.A.
 RC TISSUE=WHOLE BODY;
 RX MEDLINE=96061635; PubMed=7480137;
 RA Ohmura Y., Ohba N., Toh H., Tsuji F.I.;
 RT "Cloning, expression and sequence analysis of cDNA for the luciferases
 from the Japanese fireflies, *Pyrococcus miyako* and *Hotaria parvula*,"
 RL Photochem. Photobiol. 62:309-313(1995).

DR EMBL: U39928; AAC37254.1; -.
 DR HSSP: P08659; 1LC1.
 DR InterPro: IPR000873; AMP-bind.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP-binding; 1.
 DR PROSITE: PS00284; Serpin; UNKNOWN; 1.
 DR SEQUENCE 548 AA; 60956 MW; CD37F63E982A9401 CRC64;

Query Match 67.0%; Score 1892; DB 5; Length 548;
 Best Local Similarity 67.8%; Pred. No. 2.1e-122;
 Matches 366; Conservative 72; Mismatches 100; Indels 2; Gaps 2;

QY 4 MEND-ENIYVGEPPYPIEGSGAGOLRKYMRYAKL-GAIFATNALGVDTYAEYLEK 61
 DB 1 MEDDSKHHMGRHSILIMEDGTAGEDLHKAMRYAOVPTTFTDHAENVNITYSEYFEM 60
 QY 62 SCCLGEALKNYGLVDGRIALCSECEEFIPVLAGLFIGVAPPTNETIYRELVHSLGI 121
 DB 61 SCRLAETMKRYGLQHHIIVACSETSLQFEMPVCGALFIGVAPPTNDIYNRELYNSLI 120
 QY 122 ISKPTIVSSKKGLDVITYQKTVTAIKTIYVILDSKVDYRGYSMDNFIKKNTPOGFKGS 181
 DB 121 ISQPTIVCSKALOKILGVOKLPVIOKTVILDSREYMGQSMVSEFESHLPAGFNEY 180
 QY 182 SEKTVENRKEQVALIMNSSGSTGLPKGVOLTHENITVTRFSHARDPIYGNVSPGTALTY 241
 DB 181 DIYDPSFRETATLALIMNSSGSTGLPKGVOLTHNMVCFRSHCRDPVGNQIIPDTALIT 240
 QY 242 VYRPHHGGMTFTLYLTCGRIVMLTKFDETFELTKDYKCSSVILVPTLFAILNSEL 301
 DB 241 VYRPHHGGMTFTLYLTCGRIVMLTKFDETFELTKDYKCSSVILVPTLFAILNSEL 300
 QY 302 LLDKYDLSNLVEIASGAPLSEIGEAVARRFNLPGVROGYGLTETSALITPEDDDKP 361
 DB 301 LVDKYDLSNLHEIASGAPLSEIGEAVARRFNLPGVROGYGLTETSALITPEDDDKP 360
 QY 362 GASGVVPLFAKVIDLDTKKTLLGPNRGEVCCVGMMLKGVNDPEATREIIDEGBMLH 421
 DB 361 GASGVVPLFAKVIDLDTKKTLLGPNRGEVCCVGMMLKGVNDPEATREIIDEGBMLH 420
 QY 422 TGDGYVDEEHFIVDRKSLIKYGOVPAPELESVILQHPNIPDAGVACVPPPIAGE 481
 DB 421 SGDLATYDKDHFIVDRKSLIKYGOVPAPELESVILQHPNIPDAGVACVPPPIAGE 480
 QY 482 LPGAIVVLKKKSMTKEKEMDVYASQVSNARKLRGVRFVDEVPKGLTGKIDGKAIREIL 541
 DB 481 LPGAIVVLKKKSMTKEKEMDVYASQVSNARKLRGVRFVDEVPKGLTGKIDGKAIREIL 540

RESULT 10
 Q94697
 ID Q94697 PRELIMINARY; PRT; 552 AA.

AC Q94697;
 DT 01-FEB-1997 (TREMblrel. 02, Created)
 DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE LUCIFERASE.
 GN LOC.
 OS Photuris pennsylvanica.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Elateriformia; Cantharoidea; Lampyridae; Photuris.
 OX NCBI_TaxID=41716;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LANTERN;
 RA Zenko S., Shiratsishi S., Inouye S., Saigo K.;
 RT "Cloning, nucleotide sequence and expression of two cDNAs encoding for
 luciferase from Photuris firefly."
 RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
 DR EMBL: D25416; BAA05006.1; -.

DR HSSP: P08659; 1LC1.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP-binding; 1.
 DR SEQUENCE 552 AA; 60970 MW; CA66179AE26561DE CRC64;

Query Match 64.9%; Score 1831; DB 5; Length 552;
 Best Local Similarity 62.7%; Pred. No. 3.4e-118;
 Matches 339; Conservative 91; Mismatches 109; Indels 2; Gaps 2;

QY 4 MENDENIYVGEPPYPIEGSGAGOLRKYMRYAKL-GAIFATNALGVDTYAEYLEKS 62
 DB 1 MLEIENIILGPPPIYPIELGCTAGEDLHRAISRAAPGLTATVDTHTLEVYKEFLDYT 60
 QY 63 CCLGEALKNYGLVDGRIALCSECEEFIPVLAGLFIGVAPPTNETIYRELVHSLGI 122
 DB 61 CRLAEAMKNYGLQHTIIVSCSENCVQFMPICALYGVAPPTNDIYNRELYNSLSI 120
 QY 123 SKPTIVSSKKGLDVITYQKTVTAIKTIYVILDSKVDYRGYSMDNFIKKNTPOGFKGS 182
 DB 121 SQPTIVFTSRNSLQKILGVOSRLPIIKIILIDGKRDYLGYSQSMQFMKEHVPANFNVSA 180
 QY 183 EKTVEYNRKEQVALIMNSSGSTGLPKGVOLTHENITVTRFSHARDPIYGNVSPGTALTY 242
 DB 181 EKPPLSD-LDRVACALIMNSSGSTGLPKGVOLTHNMVCFRSHCRDPVGNQIIPDTALIT 239
 QY 243 VYRPHHGGMTFTLYLTCGRIVMLTKFDETFELTKDYKCSSVILVPTLFAILNSEL 302
 DB 240 VYRPHHGGMTFTLYLTCGRIVMLTKFDETFELTKDYKCSSVILVPTLFAILNSEL 299
 QY 303 LDKYDLSNLVEIASGAPLSEIGEAVARRFNLPGVROGYGLTETSALITPEDDDKP 362
 DB 300 VDKYDLSNLHEIASGAPLSEIGEAVARRFNLPGVROGYGLTETSALITPEDDDKP 359
 QY 363 ASGVVPLFAKVIDLDTKKTLLGPNRGEVCCVGMMLKGVNDPEATREIIDEGBMLH 422
 DB 360 AVGVVPLPYSLKVIDLDTKKTLLGPNRGEVCCVGMMLKGVNDPEATREIIDEGBMLH 419
 QY 423 GDIGYVDEEHFIVDRKSLIKYGOVPAPELESVILQHPNIPDAGVACVPPPIAGE 482
 DB 420 GDIGYVDEEHFIVDRKSLIKYGOVPAPELESVILQHPNIPDAGVACVPPPIAGE 479
 QY 483 PGAIVVLKKKSMTKEKEMDVYASQVSNARKLRGVRFVDEVPKGLTGKIDGKAIREIL 542
 DB 480 PGAIVVLKKKSMTKEKEMDVYASQVSNARKLRGVRFVDEVPKGLTGKIDGKAIREIL 539
 QY 543 K 543
 DB 540 K 540

RESULT 11
 Q94696
 ID Q94696 PRELIMINARY; PRT; 552 AA.

AC Q94696;
 DT 01-FEB-1997 (TREMblrel. 02, Created)
 DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE LUCIFERASE.
 GN LOC.
 OS Photuris pennsylvanica.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Elateriformia; Cantharoidea; Lampyridae; Photuris.
 OX NCBI_TaxID=41716;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LANTERN;
 RA Zenko S., Shiratsishi S., Inouye S., Saigo K.;
 RT "Cloning, nucleotide sequence and expression of two cDNAs encoding for
 luciferase from Photuris firefly."
 RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.

DR EMBL: D25415; BAA05005.1; -
 DR HSSP: P08659; IBA3.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP-BINDING; 1.
 SQ SEQUENCE 552 AA; 61000 MW; 85C14ED52BD5366A CRC64;

Query Match 64.6%; Score 1825; DB 5; Length 552;
 Best Local Similarity 62.5%; Pred. No. 8,9e-118;
 Matches 338; Conservative 92; Mismatches 109; Indels 2; Gaps 2;

DB 4 MENDENIVPEPPFPYPIEBSAGAOIRKYMDRYAKL-GAIAFNALGVDTYAELEKS 62
 1 MSIEENLILGPPPYPLEEGTAGEOHLRAITRYAAVGLATYDVTLELVYKEFLDYV 60
 63 CCGEALKNGLVVDGRIALSCENCEFFIPVLAGLFIGVAPATNEIYLRLELSLGI 122
 61 CRALAEAMKNGIGLQHTISVCSENCQVFMPCALYIGVAFPTNDIYNERELYSLSI 120
 DB 123 SKPTIVSSKKGLDKYTVQKTYTAIKTVIILDSKYDYGOSMDFIKKNTPOGKSS 182
 121 SQPTVYVTSRHSLOKILIGVOSRLPVIKKIIMLDTKKDYLGOSMOSKHEVPANFNVS 180
 DB 183 FKTYENRKKOVALIMNSGSGTGLPKGVOLTHENIVTRFSHADPTIYGNVSGTALIV 242
 181 FKPLSEF-LDRVACIMNSGSGTGLPKGVPISHRNTIYTRSHCHDPVYAGNOIIPDITLCA 239
 DB 243 VPFHNGFMFTTGLGYLTCGFRIVMLTKFDETEFLKTLQDYKCSSVILPFLALNSEL 302
 240 VPFHNAAGFTTNIGYIICGFHVLMYRNFNEHLFQTLQDYKCSSALIVPVLAFKLNPL 299
 DB 303 LDKYDLSNLVEIASGAPLSKEIGEAVARFNPVGVQGLTETTSALITTEBGGDKPG 362
 300 VDKYDLSNLVEIASGAPLSKEIGEAVARFNPVGVQGLTETTSALITTEBGGDKPG 359
 DB 363 ASGVVPLFKAKVIDLDTKTKTIGPNRGEVCAVGMPLMKGYVNDPEATREIIDEGLMHT 422
 360 AVGVVPLFKAKVIDLDTKTKTIGPNRGEVCAVGMPLMKGYVNDPEATREIIDEGLMHT 419
 DB 423 GDIGYDEEKEHFTVDRKSLIKYKGYVPPALESVLLQHPNIFDAGVAGVDPDLAGEL 482
 420 GDIGYDEEKEHFTVDRKSLIKYKGYVPPALESVLLQHPNIFDAGVAGVDPDLAGEL 479
 DB 483 PGAVVVLKKGSMTEKEMVYASOVSNARKLRGCVFVDEVPKGLTGKIDKAIKREILK 542
 480 PGAVVVLKKGSMTEKEMVYASOVSNARKLRGCVFVDEVPKGLTGKIDKAIKREILK 539
 DB 543 K 543
 DB 540 K 540

RESULT 12

ID 027757 PRELIMINARY; PRT; 545 AA.

AC 027757; 002653;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE LUCIFERIN 4-MONOOXYGENASE (EC 1.13.12.7) (LUCIFERASE).
 OS Photuris pennsylvanica.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Elateriformia; Cantharidae; Lampyridae; Photuris.
 OC NCBI_TaxID=41716;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LANTERN;
 RX MEDLINE=97307756; PubMed=9165098;
 RA Ye L., Buck L.M., Schaeffer H.J., Leach F.R.;
 RT "Cloning and sequencing of a cDNA for firefly luciferase from Photuris pennsylvanica.";

RL Biochim. Biophys. Acta 1339:39-52(1997).
 CC -1- FUNCTION: PRODUCES GREEN LIGHT WITH A WAVELENGTH OF 562 NM (BY
 CC SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: PHOTINUS LUCIFERIN + O(2) + ATP - OXIDIZED
 CC PHOTINUS LUCIFERIN + CO(2) + H(2)O + AMP + DIPHOSPHATE + LIGHT.
 CC -1- COFACTOR: REQUIRES A MAGNESIUM ION (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
 CC -1- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT
 CC COVALENT BINDING OF AMP TO THEIR SUBSTRATE.
 DR EMBL: U01240; AAB60897.1; -
 DR HSSP: P08659; ILC1.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP-BINDING; UNKNOWN.1.
 KW Oxidoreductase; Monooxygenase; Photoprotein; Luminescence; Magnesium;
 KW Peroxisome.
 FT SITE 543 545 MICROBODY TARGETING SIGNAL (POTENTIAL).
 SQ SEQUENCE 545 AA; 60649 MW; F0FE4E828047C26E CRC64;

Query Match 57.6%; Score 1625.5; DB 5; Length 545;
 Best Local Similarity 56.7%; Pred. No. 5.2e-104;
 Matches 308; Conservative 96; Mismatches 138; Indels 1; Gaps 1;

DB 7 DENIVYPPFPYPIEBSAGAOIRKYMDRYAKL-GAIAFNALGVDTYAELEKSCL 65
 3 DNKILGPPFPPLHDAGTAGEMFVALSRADISCIALTNAHTKENVLYEFLKLSGL 62
 DB 66 GEALKNGLVVDGRIALSCENCEFFIPVLAGLFIGVAPATNEIYLRLELSLGI 125
 63 AEFKRYGKLDNDTAVCSGNSGLQFLPLASLYLGIITAAVSDYIERELHSGIKP 122
 DB 126 TIVSSKKGLDKYTVQKTYTAIKTVIILDSKYDYGOSMDFIKKNTPOGKSSSEKT 185
 123 RIFGSKNTPQKYLANKSKLKYETIILIDNEDLGQYCLNLFNSQNDILDYKFKP 182
 DB 186 VEVNREQVALIMNSGSGTGLPKGVOLTHENIVTRFSHADPTIYGNVSGTALIV 245
 183 NSPRDDQVALVYFSSGTVGSKVMLTRKNIYARSHCKDPTFGNALPPTAILIVP 242
 DB 246 HNGFGFTTGLGYLTCGFRIVMLTKFDETEFLKTLQDYKCSSVILPFLALNSEL 305
 243 HNGFGFTTGLGYLTCGFRIVMLTKFDETEFLKTLQDYKCSSVILPFLALNSEL 302
 DB 306 YDLSNLVEIASGAPLSKEIGEAVARFNPVGVQGLTETTSALITTEBGGDKPG 365
 303 YDLSNLVEIASGAPLSKEIGEAVARFNPVGVQGLTETTSALITTEBGGDKPG 362
 DB 366 KVPVPLFKAKVIDLDTKTKTIGPNRGEVCAVGMPLMKGYVNDPEATREIIDEGLMHT 425
 363 KVPVPLFKAKVIDLDTKTKTIGPNRGEVCAVGMPLMKGYVNDPEATREIIDEGLMHT 422
 DB 426 GYDDEEKEHFTVDRKSLIKYKGYVPPALESVLLQHPNIFDAGVAGVDPDLAGEL 485
 423 AYDNDGHEFTVDRKSLIKYKGYVPPALESVLLQHPNIFDAGVAGVDPDLAGEL 482
 DB 486 VVVLKKGSMTEKEMVYASOVSNARKLRGCVFVDEVPKGLTGKIDKAIKREILK 545
 483 GVVVQTKYLINEQVONFVSSOVSTAKMLRGVYKFLDEIPKSGTKIDRVLMKREIKH 542
 DB 546 AKM 548
 DB 543 SKL 545

RESULT 13

ID 090408 PRELIMINARY; PRT; 545 AA.

AC 090408;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE LUCIFERASE.

DT 01-JUN-2001 (TEMBLrel. 17, last annotation update)
 DE CG6178 PROTEIN.
 GN CG6178.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 OX (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Abmayyan A., An H.-J., Andrews-Plannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borrova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman I.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Mosnig A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveril J.S., Zhan M., Zhan G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003746; AAF56245.1; .
 DR HSSP: P08659; ILCI.
 DR FLYBASE: Fggn0039156; CG6178.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 SQ SEQUENCE 544 AA; 59938 MW; E28B8DADCC988F3C CRC64;

Query Match 36.2%; Score 1023; DB 5; Length 544;
 Best Local Similarity 41.5%; Pred. No. 2.1e-62;
 Matches 231; Conservative 98; Mismatches 191; Indels 36; Gaps 11;

QY 9 NIVYGEPEYPIIEGSA--GAOLRKY-MDRYAKLG-AIAFTNALGTGVDTYAYEILEKSCC 64
 DB 9 NIVYGG-----PTERAQDSRLSGYIILDKYKSFGRITVLYDAVNGEYS-ASFMRKSTIV 63
 QY 65 -LGEALKNTGLVVDGIALCSENGCEEFPIPVLAGLFTGVGVADPNEITYTLRELVSIGIS 123
 DB 64 RLAYITLQKLVGVXKNDVYGLSENSVNFALAMFAGLAVGATVADPLANTYSDREVDHAINLS 123
 QY 124 KPIIVSSKKGLDKVITVYQTKVTAITIVILDS-----KVDYRGVQSMDFIKKNTPGCF 178
 DB 124 KPIIVSSKKGLDKVITVYQTKVTAITIVILDS-----KVDYRGVQSMDFIKKNTPGCF 178

DB 124 KPIIVSSKKGLDKVITVYQTKVTAITIVILDS-----KVDYRGVQSMDFIKKNTPGCF 180
 QY 179 KGSSFKTVEVNRKQVALINMSSGSTGLPKGVOLTHENIYTRSHADPIYGNVSP--- 235
 DB 181 ---DETSPAANKDSDVSLIVCSSGSGTGLPKGVOLTKNNLLATLD-----SQIOPTVI 229
 QY 236 ---GTALLTVPPFHNGFMFTTGLYLCGFRIVALTGFDEETFLKTLQDYKCSSVILVPT 292
 DB 230 PMEEVTLITVYPMFHAGCULLITTAACVGARLYLPLPEEKLFSALIEKRYMMAFVWP 289
 QY 293 LPAIALNSLDDKYDLSNLVEIASGAPLSKEIGEAVALRRFNLPGVROGYGLTETSATI 352
 DB 290 LMVFLAKHPYVDKXDLSSILWLLCGAAPLSRETDQIKERIGVFIQGYGLSESTLSVL 349
 QY 353 ITPEGDGKPGASGVVPLFAKVIDLPTKTKTIGNRRGEVCGVPMKGVNDPEATRE 412
 DB 350 VONDEFKPGSVGLKVGITYAKVIDPDTGKLGNENGELCFKDGIMKGYIGPTKSTOT 409
 QY 413 IIDECEMLHTGDIQYDEKHFIVDRLSLTKYKGYQVPAALESVILQHPNIFDAGVA 472
 DB 410 AI-KDGLHTGDIQYDDEFEFTVIDRIKELIKKGYVPAEIEALLTNDKIKDAVI 468
 QY 473 GVDPPIAGELPGAVVVLKKGKSMTEKEVMDYVASQVSNARLRGVRFPVDEPKLTGTI 532
 DB 469 GKPDDEAGELPLAFVVKQANVOLTENEVIOFVNDNASPARLRGVIFVDEIPKNSGKI 528
 QY 533 DGKAIPEILKKPVAKM 548
 DB 529 LRLIREMLKKOKSKL 544

Search completed: November 29, 2001, 01:59:33
 Job time: 378 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2001, 01:00:20 ; Search time 26.42 Seconds
(without alignments)
466,760 Million cell updates/sec

Title: US-09-581-241-4

Perfect score: 2823
Sequence: 1 MEMMENDENIVGPEPPFPI.....TKKIDKAIKRLKRVAKM 548

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 segs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents-AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTCUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2819	99.9	548	2	US-08-460-934-2
2	2819	99.9	548	2	US-08-782-118-2
3	2817	99.8	548	3	US-09-111-752-14
4	2814	99.7	548	1	US-07-903-047-8
5	2814	99.7	548	4	US-09-380-061B-16
6	2814	99.7	636	2	US-08-460-934-9
7	2814	99.7	636	2	US-08-782-118-9
8	2801	99.2	568	2	US-08-460-934-6
9	2801	99.2	568	2	US-08-782-118-6
10	2703	95.7	552	3	US-09-111-752-10
11	2684	95.1	548	1	US-07-675-211-2
12	2684	95.1	548	1	US-07-903-047-2
13	2684	95.1	548	1	US-08-076-042-2
14	2684	95.1	548	4	US-09-380-061B-14
15	2588	91.7	552	3	US-09-111-752-7
16	2394	84.8	552	3	US-09-111-752-5
17	2335.5	82.7	548	4	US-09-380-061B-18
18	1965.5	69.6	550	1	US-09-380-061B-20
19	1965.5	69.3	550	1	US-08-354-240A-4
20	1947.5	69.0	550	1	US-08-354-240A-6
21	1945.5	68.9	549	1	US-08-354-240A-2
22	1945.5	68.9	550	3	US-08-867-352-23
23	1945.5	68.9	550	4	US-09-380-061B-21
24	1945.5	68.9	815	1	US-08-122-520C-9
25	1944.5	68.9	550	4	US-08-718-425-5
26	1942.5	68.8	550	4	US-08-875-277A-4
27	1939.5	68.7	550	4	US-08-718-425-2

28	1939.5	68.7	550	4	US-08-875-277A-2	Sequence 2, Appl1
29	1937.5	68.6	550	4	US-09-380-061B-6	Sequence 6, Appl1
30	1905.5	67.5	561	2	US-08-474-169-8	Sequence 8, Appl1
31	1831	64.9	552	1	US-08-231-729B-6	Sequence 6, Appl1
32	724.5	25.7	540	4	US-08-991-677-8	Sequence 8, Appl1
33	585	20.7	578	4	US-08-981-215-1	Sequence 1, Appl1
34	258.5	9.2	798	2	US-08-222-617A-8	Sequence 8, Appl1
35	256	9.1	582	3	US-08-403-852D-20	Sequence 20, Appl1
36	256	9.1	582	3	US-08-510-646B-21	Sequence 21, Appl1
37	256	9.1	582	4	US-09-231-818-20	Sequence 20, Appl1
38	253	9.0	2618	4	US-09-413-814-28	Sequence 28, Appl1
39	251.5	8.9	669	2	US-07-861-800-2	Sequence 2, Appl1
40	251	8.9	2628	4	US-09-413-814-11	Sequence 11, Appl1
41	243	8.6	3665	2	US-08-222-617A-13	Sequence 13, Appl1
42	243	8.6	3712	2	US-08-222-617A-4	Sequence 4, Appl1
43	243	8.6	3712	2	US-08-222-617A-25	Sequence 25, Appl1
44	231	8.2	768	2	US-08-222-617A-5	Sequence 5, Appl1
45	231	8.2	3666	2	US-08-222-617A-12	Sequence 12, Appl1

ALIGNMENTS

RESULT 1
US-08-460-934-2
Sequence 2, Application US/08460934
Patent No. 5814465

GENERAL INFORMATION:
APPLICANT: TATSUMI, HIROKI
APPLICANT: KIKUCHI, SATOSHI
APPLICANT: KIKUCHI, YASUJI
APPLICANT: KIKUCHI, YASUJI
TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT
TITLE OF INVENTION: ANALYSIS METHOD
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,934
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: JP 193798/1994
FILING DATE: 27-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 54625/1995
FILING DATE: 14-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 98857/1995
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids

TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Luciola lateralis
 US-08-460-934-2

Query Match 99.98; Score 2819; DB 2; Length 548;
 Best Local Similarity 99.88; Pred. No. 3.4e-286;
 Matches 547; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNNNDENIYGGPEPPPIPIEGSAGALRKYMRYAKLGAIAFTNALTGVDYTYAYLE 60
 DB 1 MNNNDENIYGGPEPPPIPIEGSAGALRKYMRYAKLGAIAFTNALTGVDYTYAYLE 60
 QY 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFFIPVLAGLFGVGAPAPNEIYTLRELVS 120
 DB 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFFIPVLAGLFGVGAPAPNEIYTLRELVS 120
 QY 121 GISKPTIVSSKGLDVTYVQKTVAIKTIIVILDSKVDRGYQSDMNFITKKNTPGFGK 180
 DB 121 GISKPTIVSSKGLDVTYVQKTVAIKTIIVILDSKVDRGYQSDMNFITKKNTPGFGK 180
 QY 181 SSFTEVNRKEQVALLMNSGSGTLKPKVQLTNENLVTRFSHARPIYGNVSPGTAIL 240
 DB 181 SSFTEVNRKEQVALLMNSGSGTLKPKVQLTNENLVTRFSHARPIYGNVSPGTAIL 240
 QY 241 TVPFHGHGFMFTLLGLTGCGRIVMLTKFDEETFLKTLQDYKSSVILVPTLFAILNRS 300
 DB 241 TVPFHGHGFMFTLLGLTGCGRIVMLTKFDEETFLKTLQDYKSSVILVPTLFAILNRS 300
 QY 301 ELLDKYDLSNLVEIASGAPLSKEIGAVARRNLPGVRGGLTETTSAILITPBGDDK 360
 DB 301 ELLDKYDLSNLVEIASGAPLSKEIGAVARRNLPGVRGGLTETTSAILITPBGDDK 360
 QY 361 PGASGVVPLFKAKVIDLDTKTLGPNRGEVCVKPMLMKGYVNDPEATREIIDEEGWL 420
 DB 361 PGASGVVPLFKAKVIDLDTKTLGPNRGEVCVKPMLMKGYVNDPEATREIIDEEGWL 420
 QY 421 HTGDIGYDEKHFHYDRLKSLIKTKGYOVPAPLESVILLOHPNIFDAGVAGVDPPIAG 480
 DB 421 HTGDIGYDEKHFHYDRLKSLIKTKGYOVPAPLESVILLOHPNIFDAGVAGVDPPIAG 480
 QY 481 ELPGAVVVLKKGSMTEKEMDVVASOVSNARKELRGVRFVDEVPKGLTKIDGKAIREI 540
 DB 481 ELPGAVVVLKKGSMTEKEMDVVASOVSNARKELRGVRFVDEVPKGLTKIDGKAIREI 540
 QY 541 LKKPVAKM 548
 DB 541 LKKPVAKM 548

RESULT 2

US-08-782-118-2
 Sequence 2, Application US/08782118
 Patent No. 5843746

GENERAL INFORMATION:
 APPLICANT: TATSUMI, HIROKI
 APPLICANT: FUKUDA, SATOSHI
 APPLICANT: KIKUCHI, MAMORU
 APPLICANT: KOYAMA, YASUJI
 TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
 TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
 TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
 ADDRESSEE: P.C.
 STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
 CITY: ARLINGTON

STATE: VA
 COUNTRY: USA
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/782,118
 FILING DATE: 13-JAN-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/460,934
 FILING DATE: 05-JUN-1995
 APPLICATION NUMBER: JP 193798/1994
 FILING DATE: 27-JUL-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 54625/1995
 FILING DATE: 14-MAR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 98857/1995
 FILING DATE: 24-APR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: OBLON, NORMAN F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 7126-001-0
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 548 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Luciola lateralis
 US-08-782-118-2

Query Match 99.98; Score 2819; DB 2; Length 548;
 Best Local Similarity 99.88; Pred. No. 3.4e-286;
 Matches 547; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNNNDENIYGGPEPPPIPIEGSAGALRKYMRYAKLGAIAFTNALTGVDYTYAYLE 60
 DB 1 MNNNDENIYGGPEPPPIPIEGSAGALRKYMRYAKLGAIAFTNALTGVDYTYAYLE 60
 QY 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFFIPVLAGLFGVGAPAPNEIYTLRELVS 120
 DB 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFFIPVLAGLFGVGAPAPNEIYTLRELVS 120
 QY 121 GISKPTIVSSKGLDVTYVQKTVAIKTIIVILDSKVDRGYQSDMNFITKKNTPGFGK 180
 DB 121 GISKPTIVSSKGLDVTYVQKTVAIKTIIVILDSKVDRGYQSDMNFITKKNTPGFGK 180
 QY 181 SSFTEVNRKEQVALLMNSGSGTLKPKVQLTNENLVTRFSHARPIYGNVSPGTAIL 240
 DB 181 SSFTEVNRKEQVALLMNSGSGTLKPKVQLTNENLVTRFSHARPIYGNVSPGTAIL 240
 QY 241 TVPFHGHGFMFTLLGLTGCGRIVMLTKFDEETFLKTLQDYKSSVILVPTLFAILNRS 300
 DB 241 TVPFHGHGFMFTLLGLTGCGRIVMLTKFDEETFLKTLQDYKSSVILVPTLFAILNRS 300
 QY 301 ELLDKYDLSNLVEIASGAPLSKEIGAVARRNLPGVRGGLTETTSAILITPBGDDK 360
 DB 301 ELLDKYDLSNLVEIASGAPLSKEIGAVARRNLPGVRGGLTETTSAILITPBGDDK 360
 QY 361 PGASGVVPLFKAKVIDLDTKTLGPNRGEVCVKPMLMKGYVNDPEATREIIDEEGWL 420
 DB 361 PGASGVVPLFKAKVIDLDTKTLGPNRGEVCVKPMLMKGYVNDPEATREIIDEEGWL 420

QY 421 HTGDIGYDEEKEHFIYDRLSLIKYGYVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
DB 421 HTGDIGYDEEKEHFIYDRLSLIKYGYVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
QY 481 ELPGAVVVLKKGKSMTEKEVMDYVASQVSNAKRLRGVRFVDEVPKGLTGKIDKAIKEI 540
DB 481 ELPGAVVVLKKGKSMTEKEVMDYVASQVSNAKRLRGVRFVDEVPKGLTGKIDKAIKEI 540
QY 541 LKRPVAKM 548
DB 541 LKRPVAKM 548

RESULT 3
US-09-111-752-14
Sequence 14, Application US/09111752
Patent No. 6074859
GENERAL INFORMATION:
APPLICANT: HITOKAWA, KOZO
APPLICANT: KAJIYAMA, NAOKI
APPLICANT: MORAKAMI, SEIJI
TITLE OF INVENTION: MUTANT-TYPE BIOLUMINESCENT PROTEIN, AND
TITLE OF INVENTION: PROCESS FOR PRODUCING MUTANT-TYPE LUMINESCENT PROTEIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/111,752
FILING DATE: 08-JUL-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-0009-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-111-752-14

Query Match 99.8%; Score 2817; DB 3; Length 548;
Best Local Similarity 99.8%; Pred. No. 5.5e-286;
Matches 547; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MENMENDENIVYGPPEPIEGSGAGOLRKYMRYAKIGATATFNALTGVDYTYAEYLE 60
DB 1 MENMENDENIVYGPPEPIEGSGAGOLRKYMRYAKIGATATFNALTGVDYTYAEYLE 60
QY 61 KSCGCEALKNGVLVDGRIALCSCNCEFFIPVLGLFTIGVAPATNEIYTLRELIVSL 120
DB 61 KSCGCEALKNGVLVDGRIALCSCNCEFFIPVLGLFTIGVAPATNEIYTLRELIVSL 120
QY 121 GISKPTIVSSKKGLDKVITVOKTVAITKIYILDSKDYRGYSMDNFIRKNTQGRKG 180
DB 121 GISKPTIVSSKKGLDKVITVOKTVAITKIYILDSKDYRGYSMDNFIRKNTQGRKG 180

DB 121 GISKPTIVSSKKGLDKVITVOKTVAITKIYILDSKDYRGYSMDNFIRKNTQGRKG 180
QY 181 SSFKTYEVRNKRGOVALINSSGSTGLPKGVOLTHENLVTRFSHARDPIYGNVSPGTAIL 240
DB 181 SSFKTYEVRNKRGOVALINSSGSTGLPKGVOLTHENLVTRFSHARDPIYGNVSPGTAIL 240
QY 241 TVVFFHNGFGMFTLLGYLTCGFRIVMLTKFDETEFLKTLQDYKCSSVILVPTLAILNRS 300
DB 241 TVVFFHNGFGMFTLLGYLTCGFRIVMLTKFDETEFLKTLQDYKCSSVILVPTLAILNRS 300
QY 301 ELLDKYDLSNLVPIASGAPLSEIGEAVARRNPLPGVROGYGLTETTSAILITPEGDOK 360
DB 301 ELLDKYDLSNLVPIASGAPLSEIGEAVARRNPLPGVROGYGLTETTSAILITPEGDOK 360
QY 361 PGASGVVPLFKAKVIDLTFTKTLGPNRRCGECVKGPMIMKGVVDNPEATREIIDEGBGL 420
DB 361 PGASGVVPLFKAKVIDLTFTKTLGPNRRCGECVKGPMIMKGVVDNPEATREIIDEGBGL 420
QY 421 HTGDIGYDEEKEHFIYDRLSLIKYGYVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
DB 421 HTGDIGYDEEKEHFIYDRLSLIKYGYVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
QY 481 ELPGAVVVLKKGKSMTEKEVMDYVASQVSNAKRLRGVRFVDEVPKGLTGKIDKAIKEI 540
DB 481 ELPGAVVVLKKGKSMTEKEVMDYVASQVSNAKRLRGVRFVDEVPKGLTGKIDKAIKEI 540
QY 541 LKRPVAKM 548
DB 541 LKRPVAKM 548

RESULT 4
US-07-903-047-8
Sequence 8, Application US/07903047
Patent No. 5229285
GENERAL INFORMATION:
APPLICANT: KajiYama, Naoki
APPLICANT: Nakano, Eiichi
TITLE OF INVENTION: Thermostable luciferase Of Firefly,
TITLE OF INVENTION: Thermostable luciferase Gene Of Firefly, No. 5229285a1 Reco
TITLE OF INVENTION: DNA, And Process For The Preparation Of Thermostable
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Penile & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/903,047
FILING DATE: 19920623
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7005-048
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 869-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENITE
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: unknown

Patent No. 5814465
GENERAL INFORMATION:
APPLICANT: TATSUMI, HIROKI
APPLICANT: FUKUDA, SATOSHI
APPLICANT: KIKUCHI, MAMORU
APPLICANT: KOYAMA, YASUJI
TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINATED AND A BIOLUMINESCENT
TITLE OF INVENTION: ANALYSIS METHOD
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,934
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 193798/1994
FILING DATE: 27-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 54625/1995
FILING DATE: 14-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 98857/1995
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ. ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 636 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-934-9

Query Match 99.7%; Score 2814; DB 2; Length 636;
Best Local Similarity 99.8%; Pred. No. 1.5e-285;
Matches 546; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEMMENDENIVGPEPEFYIEBSSAGAOIRKYMWRKAKIGALAFNALGVDTTAEYLE 60
DB 1 MEMMENDENIVGPEPEFYIEBSSAGAOIRKYMWRKAKIGALAFNALGVDTTAEYLE 60
QY 61 KSCCLGALKNGVLVDGRITALSCENCEFEFIPVLGLFVGAVAPNTNITLRELHSL 120
DB 61 KSCCLGALKNGVLVDGRITALSCENCEFEFIPVLGLFVGAVAPNTNITLRELHSL 120
QY 121 GISKPTIVSSKKGDKLVITVQKTVAITVILDSKVDRGYQSDNFIKKNTPGQFNG 180
DB 121 GISKPTIVSSKKGDKLVITVQKTVAITVILDSKVDRGYQSDNFIKKNTPGQFNG 180
QY 181 SSFRTVEVNKKEQVALIMNMSGSTGLPKGYOLTHENLVTFRSHARDPIIGNOVSPETAIL 240
DB 181 SSFRTVEVNKKEQVALIMNMSGSTGLPKGYOLTHENLVTFRSHARDPIIGNOVSPETAIL 240

QY 241 TVPFFHHGFMFTTGLYTCGFRIVMLTKFDEETFLTKLDYKCSSYILVPTLFAILNRS 300
DB 241 TVPFFHHGFMFTTGLYTCGFRIVMLTKFDEETFLTKLDYKCSSYILVPTLFAILNRS 300
QY 301 ELLDKYDLSNLVEIASGAPLSKEIGAVARRENLPGVROGYGLTETTSAILITTPGDDK 360
DB 301 ELLDKYDLSNLVEIASGAPLSKEIGAVARRENLPGVROGYGLTETTSAILITTPGDDK 360
QY 361 PGASGVVPLFKAVIDLDTKTIGPBRGECVCKGMLKGYVDNPEATREIIDEGBWL 420
DB 361 PGASGVVPLFKAVIDLDTKTIGPBRGECVCKGMLKGYVDNPEATREIIDEGBWL 420
QY 421 HTGDIGYDEKHFPIYDRLSLIKRYGOVPPALESVLLQHPNIEDACVAGVPPPIAG 480
DB 421 HTGDIGYDEKHFPIYDRLSLIKRYGOVPPALESVLLQHPNIEDACVAGVPPPIAG 480
QY 481 ELPGAIVVLLKKGKSMTEKEVMDIYASOVSNARKLRGVRVDEVPKGLTKIDGKAIREI 540
DB 481 ELPGAIVVLLKKGKSMTEKEVMDIYASOVSNARKLRGVRVDEVPKGLTKIDGKAIREI 540
QY 541 LKKPVAK 547
DB 541 LKKPVAK 547

RESULT 7
US-08-782-118-9
Sequence 9, Application US/08782118
Patent No. 5843746
GENERAL INFORMATION:
APPLICANT: TATSUMI, HIROKI
APPLICANT: FUKUDA, SATOSHI
APPLICANT: KIKUCHI, MAMORU
APPLICANT: KOYAMA, YASUJI
TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINATED AND A BIOLUMINESCENT
TITLE OF INVENTION: ANALYSIS METHOD
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/782,118
FILING DATE: 13-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/460,934
FILING DATE: 05-JUN-1995
APPLICATION NUMBER: JP 193798/1994
FILING DATE: 27-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 54625/1995
FILING DATE: 14-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 98857/1995
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000

TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 636 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-782-118-9

Query Match 99.7%; Score 2814; DB 2; Length 636;
 Best Local Similarity 99.8%; Pred. No. 1.5e-285;
 Matches 546; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MENNENDENIYVGEPEFYPIIEGSGAQLRKRYMDRYAKLGAIAFTNALTGVDYTAETLE 60
 DB 1 MENNENDENIYVGEPEFYPIIEGSGAQLRKRYMDRYAKLGAIAFTNALTGVDYTAETLE 60
 QY 61 KSCCLGALKNYGLVVGRIALCSENCPEEPIPLAGLFTIGVGAAPINEIYTLRELVHSL 120
 DB 61 KSCCLGALKNYGLVVGRIALCSENCPEEPIPLAGLFTIGVGAAPINEIYTLRELVHSL 120
 QY 121 GISKPTIVSSKGLDVTYVOKTVAIKTIVILDSKYDYGYSMDNEFIKKNPQGFKG 180
 DB 121 GISKPTIVSSKGLDVTYVOKTVAIKTIVILDSKYDYGYSMDNEFIKKNPQGFKG 180
 QY 181 SSKFTVEVNRKEOVALLIMNSGSTGLPKGVOLTHEMLVTRFSHARDPIYGNVSPGTAIL 240
 DB 181 SSKFTVEVNRKEOVALLIMNSGSTGLPKGVOLTHEMLVTRFSHARDPIYGNVSPGTAIL 240
 QY 241 TYVPHHGFMFTTLGLTGGRIVMLTKPDEEFLTKTDYKSSVILVPTLAILNRS 300
 DB 241 TYVPHHGFMFTTLGLTGGRIVMLTKPDEEFLTKTDYKSSVILVPTLAILNRS 300
 QY 301 ELLDKYDLSNLVEIASGAPLSKEIGAVARFNLPGVROGYGLTETTSAILIIPEDGDK 360
 DB 301 ELLDKYDLSNLVEIASGAPLSKEIGAVARFNLPGVROGYGLTETTSAILIIPEDGDK 360
 QY 361 PGASGRVPLFKAKVIDLDRKTKLTPNRGECVCKGPMKGYVDNEAREIIDEESGWL 420
 DB 361 PGASGRVPLFKAKVIDLDRKTKLTPNRGECVCKGPMKGYVDNEAREIIDEESGWL 420
 QY 421 HTDDIGVDEEKFHTYDRKSLIKYGYOVPALESVLLQHPNIFDAGVAGVDDIAG 480
 DB 421 HTDDIGVDEEKFHTYDRKSLIKYGYOVPALESVLLQHPNIFDAGVAGVDDIAG 480
 QY 481 ELPGAVVILKKGSMTEKEVMDYVAVSQVSNAKRLRGVREDEVPKGLTKIDGKAIREI 540
 DB 481 ELPGAVVILKKGSMTEKEVMDYVAVSQVSNAKRLRGVREDEVPKGLTKIDGKAIREI 540
 QY 541 LKRPVAK 547
 DB 541 LKRPVAK 547

RESULT 8
 US-08-460-934-6
 Sequence 6, Application US/08460934
 Patent No. 5814465
 GENERAL INFORMATION:
 APPLICANT: TATSUMI, HIROKI
 APPLICANT: FUKUDA, SATOSHI
 APPLICANT: KIKUCHI, MAMORU
 APPLICANT: KOTAMA, YASUHI
 TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
 TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
 TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
 ADDRESSEE: P.C.
 STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR

CITY: ARLINGTON
 STATE: VA
 COUNTRY: USA
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/460,934
 FILING DATE: 05-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 193798/1994
 FILING DATE: 27-JUL-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 54625/1995
 FILING DATE: 14-MAR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 98857/1995
 FILING DATE: 24-APR-1995
 AUTHOR/AGENT INFORMATION:
 NAME: OBLON, NORMAN F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 7126-001-0
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 568 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-460-934-6

Query Match 99.2%; Score 2801; DB 2; Length 568;
 Best Local Similarity 99.5%; Pred. No. 2.8e-284;
 Matches 543; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 MNENDENIYVGEPEFYPIIEGSGAQLRKRYMDRYAKLGAIAFTNALTGVDYTAETLE 62
 DB 23 SLENDENIYVGEPEFYPIIEGSGAQLRKRYMDRYAKLGAIAFTNALTGVDYTAETLE 62
 QY 63 CCLGALKNYGLVVGRIALCSENCPEEPIPLAGLFTIGVGAAPINEIYTLRELVHSL 122
 DB 63 CCLGALKNYGLVVGRIALCSENCPEEPIPLAGLFTIGVGAAPINEIYTLRELVHSL 122
 QY 123 GISKPTIVSSKGLDVTYVOKTVAIKTIVILDSKYDYGYSMDNEFIKKNPQGFKG 182
 DB 123 GISKPTIVSSKGLDVTYVOKTVAIKTIVILDSKYDYGYSMDNEFIKKNPQGFKG 182
 QY 183 SSKFTVEVNRKEOVALLIMNSGSTGLPKGVOLTHEMLVTRFSHARDPIYGNVSPGTAIL 242
 DB 183 SSKFTVEVNRKEOVALLIMNSGSTGLPKGVOLTHEMLVTRFSHARDPIYGNVSPGTAIL 242
 QY 241 TYVPHHGFMFTTLGLTGGRIVMLTKPDEEFLTKTDYKSSVILVPTLAILNRS 302
 DB 241 TYVPHHGFMFTTLGLTGGRIVMLTKPDEEFLTKTDYKSSVILVPTLAILNRS 302
 QY 301 ELLDKYDLSNLVEIASGAPLSKEIGAVARFNLPGVROGYGLTETTSAILIIPEDGDK 362
 DB 301 ELLDKYDLSNLVEIASGAPLSKEIGAVARFNLPGVROGYGLTETTSAILIIPEDGDK 362
 QY 361 PGASGRVPLFKAKVIDLDRKTKLTPNRGECVCKGPMKGYVDNEAREIIDEESGWL 422
 DB 361 PGASGRVPLFKAKVIDLDRKTKLTPNRGECVCKGPMKGYVDNEAREIIDEESGWL 422
 QY 421 HTDDIGVDEEKFHTYDRKSLIKYGYOVPALESVLLQHPNIFDAGVAGVDDIAG 482
 DB 421 HTDDIGVDEEKFHTYDRKSLIKYGYOVPALESVLLQHPNIFDAGVAGVDDIAG 482
 QY 443 GDIGYDEEKFHTYDRKSLIKYGYOVPALESVLLQHPNIFDAGVAGVDDIAG 502
 DB 443 GDIGYDEEKFHTYDRKSLIKYGYOVPALESVLLQHPNIFDAGVAGVDDIAG 502

QY 483 PGAVVVLKKGKSMTEKEVMDYVAVSOVSNARKLRGVRVDEVPKGLTGKIDGKAIKREILK 542
DB 503 PGAVVVLKKGKSMTEKEVMDYVAVSOVSNARKLRGVRVDEVPKGLTGKIDGKAIKREILK 562
QY 543 KPVAKM 548
DB 563 KPVAKM 568

RESULT 9

US-08-782-118-6
; Sequence 6, Application US/08782118
; Patent No. 5843746
GENERAL INFORMATION:
APPLICANT: TATSUMI, HIROKI
APPLICANT: FUKUDA, SATOSHI
APPLICANT: KIKUCHI, MAMORU
APPLICANT: KOYAMA, YASUJI
TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/782,118
FILING DATE: 13-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/460,934
FILING DATE: 05-JUN-1995
APPLICATION NUMBER: JP 193798/1994
FILING DATE: 27-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 54625/1995
FILING DATE: 14-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 98857/1995
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 568 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-782-118-6

Query Match 99.2%; Score 2801; DB 2; Length 568;
Best Local Similarity 99.5%; Pred. No. 2,8e-284;
Matches 543; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 3 NMENDENIVGPEPFYPIEESAGAOQLRKYMRYAKIGAIATFNALGVDTYAEYLEKS 62

DB 23 SLENDENIVGPEPFYPIEESAGAOQLRKYMRYAKIGAIATFNALGVDTYAEYLEKS 82
QY 63 CCEBALKNGLVVDGRALALSCENCEEFFIPVLAGLFIGVAPATNTIYLRRLVSLGI 122
DB 83 CCEBALKNGLVVDGRALALSCENCEEFFIPVLAGLFIGVAPATNTIYLRRLVSLGI 142
QY 123 SKPTIVSSKKGLDKVITVOKTVAITKIYILDSKYDVRGYSMDNFIKNTPOGFSS 182
DB 143 SKPTIVSSKKGLDKVITVOKTVAITKIYILDSKYDVRGYSMDNFIKNTPOGFSS 202
QY 183 FKTEVNRKEQVALINMSGSGTLPGKVOLEHNLVTRFSHARDPIYGNVSPGATILTV 242
DB 203 FKTEVNRKEQVALINMSGSGTLPGKVOLEHNLVTRFSHARDPIYGNVSPGATILTV 262
QY 243 VPFHFGMFTTLGTLGCGFRIWMLTFDEFTPLTKIDYKCSSVILVPLFALINSEL 302
DB 263 VPFHFGMFTTLGTLGCGFRIWMLTFDEFTPLTKIDYKCSSVILVPLFALINSEL 322
QY 303 LDKYDLSNLVEIASGAPLSKEIGEAVARPNLPGRVGGYGLFTTSAIITPEGDGPG 362
DB 323 LDKYDLSNLVEIASGAPLSKEIGEAVARPNLPGRVGGYGLFTTSAIITPEGDGPG 382
QY 363 ASGKVPFLKAKVIDLTKTGLGNRRGECVCKPMLDKGYDNPETREIIDEGLMT 422
DB 383 ASGKVPFLKAKVIDLTKTGLGNRRGECVCKPMLDKGYDNPETREIIDEGLMT 442
QY 423 GDIGYDEKHFPIYDRKSLIKKYGQVPALESLVLOHPNIFDGVAGVPPINGEL 482
DB 443 GDIGYDEKHFPIYDRKSLIKKYGQVPALESLVLOHPNIFDGVAGVPPINGEL 502
QY 483 PGAVVVLKKGKSMTEKEVMDYVAVSOVSNARKLRGVRVDEVPKGLTGKIDGKAIKREILK 542
DB 503 PGAVVVLKKGKSMTEKEVMDYVAVSOVSNARKLRGVRVDEVPKGLTGKIDGKAIKREILK 562
QY 543 KPVAKM 548
DB 563 KPVAKM 568

RESULT 10

US-09-111-752-10
; Sequence 10, Application US/09111752
; Patent No. 6074859
GENERAL INFORMATION:
APPLICANT: HIROKAWA, KOZO
APPLICANT: KAJIYAMA, NAOKI
APPLICANT: MORAKAMI, SEIJI
TITLE OF INVENTION: MUTANT-TYPE BIOLUMINESCENT PROTEIN, AND
TITLE OF INVENTION: PROCESS FOR PRODUCING MUTANT-TYPE BIOLUMINESCENT PROTEIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/111,752
FILING DATE: 08-JUL-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-0009-0
TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 552 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Luciola lateralis, Photinus pyralis
 US-09-111-752-10

Query Match 95.7%; Score 2703; DB 3; Length 552;
 Best Local Similarity 96.3%; Pred. No. 4.8e-274;
 Matches 523; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

QY 1 MENNENDENIYGEPEPPYPIEGSAGAOQKRYMDRYAKLGAIAFTNALTGVDYVAYEYL 60
 DB 1 MENNENDENIYGEPEPPYPIEGSAGAOQKRYMDRYAKLGAIAFTNALTGVDYVAYEYL 60
 QY 61 KSCCLGALKNYGLVNGRIALCSENCEEPFIVLAGLFGVGAAPNEIYTLRELHVS 120
 DB 61 KSCCLGALKNYGLVNGRIALCSENCEEPFIVLAGLFGVGAAPNEIYTLRELHVS 120
 QY 121 GISKPTIVFSSKGLDKVITVQKTVTAIKTIVILDSKVDRGYOSMDFIKKNTPOGFG 180
 DB 121 GISKPTIVFSSKGLDKVITVQKTVTAIKTIVILDSKVDRGYOSMDFIKKNTPOGFG 180
 QY 181 SSFTYVNRKEQVALIMNSSGSTGLPKGYQLHENVLTPRSHARDPIYGNVSPGTAIL 240
 DB 181 SSFTYVNRKEQVALIMNSSGSTGLPKGYQLHENVLTPRSHARDPIYGNVSPGTAIL 240
 QY 241 TVPFHNGFMFTLLGLTGFRIYMLTKFDEETFLKTLQDYKSSVILVPTLFAILNRS 300
 DB 241 TVPFHNGFMFTLLGLTGFRIYMLTKFDEETFLKTLQDYKSSVILVPTLFAILNRS 300
 QY 301 ELLDKYDLSMLVEIASGAPLSKEIGAVARFNLPGVRGYGLTETTSAILITPEGDDK 360
 DB 301 ELLDKYDLSMLVEIASGAPLSKEIGAVARFNLPGVRGYGLTETTSAILITPEGDDK 360
 QY 361 PGASGVKVPLEKAKVIDLDLTKTLGPRRGEVCKGPMKKGVDNDEARREITIDEGWL 420
 DB 361 PGASGVKVPLEKAKVIDLDLTKTLGPRRGEVCKGPMKKGVDNDEARREITIDEGWL 420
 QY 421 HTGDIGYDEEKHFIVDRLSLITKYGQVPALESVLLQHPNIFDAGVAGVPPDIAG 480
 DB 421 HTGDIGYDEEKHFIVDRLSLITKYGQVPALESVLLQHPNIFDAGVAGVPPDIAG 480
 QY 481 ELPGAAYVILKKGSMTEKEMDYVASOVNAKRLRGVRFVDEVPKGLTGIDGAIAREI 540
 DB 481 ELPGAAYVILKKGSMTEKEMDYVASOVNAKRLRGVRFVDEVPKGLTGIDGAIAREI 540
 QY 541 LKK 543
 DB 541 LKK 543

RESULT 11
 US-07-675-211-2
 ; Sequence 2, Application US/07675211
 ; Patent No. 5219737
 ; GENERAL INFORMATION:
 ; APPLICANT: KAITIYAMA, MOKI
 ; APPLICANT: NAKANO, EICHI
 ; TITLE OF INVENTION: MUTANT LUCIFERASE OF A FIREFLY, MUTANT
 ; TITLE OF INVENTION: LUCIFERASE GENES, NOVEL RECOMBINANT DNAs CONTAINING THE
 ; TITLE OF INVENTION: GENES AND A METHOD OF PRODUCING MUTANT LUCIFERASE
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PENNIE & EDMONDS
 ; STREET: 1155 AVENUE OF THE AMERICAS

CITY: NEW YORK
 STATE: N.Y.
 COUNTRY: U.S.A
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/675,211
 FILING DATE: 19910326
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MISROCK, S. LESLIE
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 7005-026-999
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-790-9090
 TELEFAX: 212-869-9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 548 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Luciola cruciata
 US-07-675-211-2

Query Match 95.1%; Score 2684; DB 1; Length 548;
 Best Local Similarity 93.4%; Pred. No. 4.6e-272;
 Matches 512; Conservative 26; Mismatches 10; Indels 0; Gaps 0;

QY 1 MENNENDENIYGEPEPPYPIEGSAGAOQKRYMDRYAKLGAIAFTNALTGVDYVAYEYL 60
 DB 1 MENNENDENIYGEPEPPYPIEGSAGAOQKRYMDRYAKLGAIAFTNALTGVDYVAYEYL 60
 QY 61 KSCCLGALKNYGLVNGRIALCSENCEEPFIVLAGLFGVGAAPNEIYTLRELHVS 120
 DB 61 KSCCLGALKNYGLVNGRIALCSENCEEPFIVLAGLFGVGAAPNEIYTLRELHVS 120
 QY 121 GISKPTIVFSSKGLDKVITVQKTVTAIKTIVILDSKVDRGYOSMDFIKKNTPOGFG 180
 DB 121 GISKPTIVFSSKGLDKVITVQKTVTAIKTIVILDSKVDRGYOSMDFIKKNTPOGFG 180
 QY 181 SSFTYVNRKEQVALIMNSSGSTGLPKGYQLHENVLTPRSHARDPIYGNVSPGTAIL 240
 DB 181 SSFTYVNRKEQVALIMNSSGSTGLPKGYQLHENVLTPRSHARDPIYGNVSPGTAIL 240
 QY 241 TVPFHNGFMFTLLGLTGFRIYMLTKFDEETFLKTLQDYKSSVILVPTLFAILNRS 300
 DB 241 TVPFHNGFMFTLLGLTGFRIYMLTKFDEETFLKTLQDYKSSVILVPTLFAILNRS 300
 QY 301 ELLDKYDLSMLVEIASGAPLSKEIGAVARFNLPGVRGYGLTETTSAILITPEGDDK 360
 DB 301 ELLDKYDLSMLVEIASGAPLSKEIGAVARFNLPGVRGYGLTETTSAILITPEGDDK 360
 QY 361 PGASGVKVPLEKAKVIDLDLTKTLGPRRGEVCKGPMKKGVDNDEARREITIDEGWL 420
 DB 361 PGASGVKVPLEKAKVIDLDLTKTLGPRRGEVCKGPMKKGVDNDEARREITIDEGWL 420
 QY 421 HTGDIGYDEEKHFIVDRLSLITKYGQVPALESVLLQHPNIFDAGVAGVPPDIAG 480
 DB 421 HTGDIGYDEEKHFIVDRLSLITKYGQVPALESVLLQHPNIFDAGVAGVPPDIAG 480
 QY 481 ELPGAAYVILKKGSMTEKEMDYVASOVNAKRLRGVRFVDEVPKGLTGIDGAIAREI 540
 DB 481 ELPGAAYVILKKGSMTEKEMDYVASOVNAKRLRGVRFVDEVPKGLTGIDGAIAREI 540
 QY 541 LKKPVAKM 548

Db 541 LKPYAKM 548

RESULT 12

US-07-903-047-2
; Sequence 2, Application US/07903047
; Patent No. 5229285
; GENERAL INFORMATION:
; APPLICANT: Kajiya, Naoki
; APPLICANT: Nakano, Eiichi
; TITLE OF INVENTION: Thermostable Luciferase Gene of Firefly,
; TITLE OF INVENTION: Thermostable Luciferase Gene of Firefly, No. 5229285el Recombi
; TITLE OF INVENTION: DNA, And Process For The Preparation Of Thermostable
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/903,047
; FILING DATE: 19920623
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7005-048
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-07-903-047-2

Query Match 95.1%; Score 2684; DB 1; Length 548;
Best Local Similarity 93.4%; Pred. No. 4 6e-272;
Matches 512; Conservative 26; Mismatches 10; Indels 0; Gaps 0;

QY 1 MEMNENDENIVGPPEFYPIEESGAGOLKRYMDRYAKLGAIAFTNALTGVDTVAEYLE 60
DB 1 MEMNENDENIVGPPEFYPIEESGAGOLKRYMDRYAKLGAIAFTNALTGVDTVAEYLE 60
QY 61 KSCCLGALKNYGLVVDGRALSCENCEFFIVLAGLPIGVAVATNNEITYTLRELVAHSI 120
DB 61 KSCCLGALKNYGLVVDGRALSCENCEFFIVLAGLPIGVAVATNNEITYTLRELVAHSI 120
QY 121 GISKPTIVFSKSGKLDKVTIVQKTVTAIKITVILDSKVDYRGVQSDMNFIKNTPOGFGK 180
DB 121 GISKPTIVFSKSGKLDKVTIVQKTVTAIKITVILDSKVDYRGVQSDMNFIKNTPOGFGK 180
QY 181 SSFRTVEVNRKEQVALIMNSGSGTGLPKGVOLTHENLVTRSHARDPIYGNVSPCTAVI 240
DB 181 SSFRTVEVNRKEQVALIMNSGSGTGLPKGVOLTHENLVTRSHARDPIYGNVSPCTAVI 240
QY 241 TVVPEHHGFMFTTGLGILGFRVYMLTKPDEFTPLTKLQDYKCSVILVPTLFAILNKS 300
DB 241 TVVPEHHGFMFTTGLGILGFRVYMLTKPDEFTPLTKLQDYKCSVILVPTLFAILNKS 300

QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRENLPGVROGYGLTETTSAILITTPGDDK 360
DB 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRENLPGVROGYGLTETTSAILITTPGDDK 360
QY 361 PGASGKVPLEFKAKVIDLDTRKTLGPNRRGEVCYKGMPLKKGVDNDEATREIIDEKGWL 420
DB 361 PGASGKVPLEFKAKVIDLDTRKTLGPNRRGEVCYKGMPLKKGVDNDEATREIIDEKGWL 420
QY 421 HTGDIGYDEEKHEFFIVDRKLSLIKRYGYVPAELESYLLQHPNIFDAGVAGPPPIAG 480
DB 421 HTGDIGYDEEKHEFFIVDRKLSLIKRYGYVPAELESYLLQHPNIFDAGVAGPPPIAG 480
QY 481 ELPGAVVLLKKGKSMTEKEVMDYVASOVNAKRLRGVRFVDEPKGLTGKIDGKATREI 540
DB 481 ELPGAVVLLKKGKSMTEKEVMDYVASOVNAKRLRGVRFVDEPKGLTGKIDGKATREI 540
QY 541 LKPYAKM 548
DB 541 LKPYAKM 548

RESULT 13

US-08-076-042-2
; Sequence 2, Application US/08076042
; Patent No. 5330906
; GENERAL INFORMATION:
; APPLICANT: KAJIYAMA, NAOKI
; APPLICANT: NAKANO, EIICHI
; TITLE OF INVENTION: MUTANT LUCIFERASE OF A FIREFLY, MUTANT
; TITLE OF INVENTION: LUCIFERASE GENES, NOVEL RECOMBINANT DNAs CONTAINING THE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: N.Y.
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/076,042
; FILING DATE: 15-JUN-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/675,211
; FILING DATE: 26-MAR-1991
; NAME: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MISTROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7005-026-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Luciola cruciata
; US-08-076-042-2

Query Match 95.1%; Score 2684; DB 1; Length 548;

Best Local Similarity 93.4%; Pred. No. 4,6e-272;
Matches 512; Conservative 26; Mismatches 10; Indels 0; Gaps 0

```

QY 1 MENMENDNIYAPPEPFIIEGSGAOLRYKMPRYAKIGIAETNNLTGVDYVAYLE 60
Dp 1 MENMENDNIYAGRPFIPIIEGSGAOLRYKMYRYAKIGIAETNNLTGVDISAYYLE 60
QY 61 KSCCLGEALKNYGLVODGRIALCSENCEEFPIPIVAGFIQGVAPANEIYTLREIYHSL 120
Dp 61 KSCCLGEALKNYGLVODGRIALCSENCEEFPIPIVAGFIQGVAPANEIYTLREIYHSL 120
QY 121 GISKPTIYESSKKGIDKITYOKVTYVTAKTIVIIIDSKVDYRGQSMNFIKKNTPQSGFK 180
Dp 121 GISKPTIYESSKKGIDKITYOKVTYVTAKTIVIIIDSKVDYRGQSOCLDFIKRNTPPGFOA 180
QY 181 SSFKTVEVNRREQVALLINSSSGTGLEKRGVOLTENIYVTFRSHARDEITYGNQVSPGAIL 240
Dp 181 SSFKTVEVNRREQVALLINSSSGTGLEKRGVOLTENIYVTFRSHARDEITYGNQVSPGAVL 240

```

RESULT 14
US-09-380-061B-14
Sequence 14, Application US/09380061B
Patent No. 6265177
GENERAL INFORMATION:
APPLICANT: SQUIRRELL, DAVID JAMES
WHITE, PETER JOHN
LOWE, CHRISTOPHER ROBIN
MURRAY, JAMES AUGUSTUS HENRY
TITLE OF INVENTION: ENZYME ASSAY FOR MUTANT FIREFLY LUCIFERASE
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLOBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/380,061B
FILING DATE: 25-Aug-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB98/01026

? FILING DATE: 7-APR-1998
 ? APPLICATION NUMBER: GB 9707468.8
 ? FILING DATE: 11-APR-1997
 ? ATTORNEY/AGENT INFORMATION:
 ? NAME: SMOOF, B. J.
 ? REGISTRATION NUMBER: 36,663
 ? REFERENCE/DOCKET NUMBER: 124-7255
 ? TELECOMMUNICATION INFORMATION:
 ? TELEPHONE: (703)816-4000
 ? TELEFAX: (703)816-4100
 ? INFORMATION FOR SEQ ID NO: 14:
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 548 amino acids
 ? TYPE: amino acid
 ? TOPOLOGY: linear
 ? MOLECULE TYPE: protein
 ? SEQUENCE DESCRIPTION: SEQ ID NO: 14:
 ? IS-09-380-061B-14

Query Match	95.1%	Score 2684	DB 4	Length 548
Best Local Similarity	93.4%	Pred. No.	4.6e-272	
Matches 512	Conservative 26	Mismatches 10	Indels 0	Gaps 0

Qy	1	MENNENENIYVGEPEFYPIEESAGAQOLKRYMYRVAKLGIAIATNALTGVDYVAYTLE	60
Db	1	MENNENDENIYVGRKPFPIEBSAGQOLKRYMYRVAKLGIAIATNAVTGVDYSAYTLE	60
Qy	61	KSCCLGEALKNYGLVVDGRJALCSENGEEFFIPVLAGJLTGVGVAPTNEIYTLREIYVSL	120
Db	61	KSCCLGALQWYGVJVDGRJALCSENGEEFFIPVJAGJLTGVGVAPTNEIYTLREIYVSL	120
Qy	121	GISKPTIVSESSKGLDVITVQKRYVTAIKTIVILDSKVODRGQSMNFIKKNTPQSFKG	180
Db	121	GISKPTIVSESSKGLDVITVQKRYVTAIKTIVILDSKVODRGYQCLDTFIKKNTPQSFQA	180
Qy	181	SSFEFTVEVNRKEOVALIMNSGSGTGLERGVOLTHEENITVFSHARDEITYNOVSPGTAIL	240
Db	181	SSFEFTVEVNRKEOVALIMNSGSGTGLERGVOLTHEENITVFSHARDEITYNOVSPGTAIVL	240
Qy	241	TVVPEFHGFGMFTTLYGLYJCGFRIVMJTKPDEFTFKTLTDYKCSSVILVPTLFAILNRS	300
Db	241	TVVPEFHGFGMFTTLYGLYJCGFRIVMJTKPDEFTFKTLTDYKCTSVILVPTLFAILNKS	300
Qy	301	ELLDKLYDLSNVLVJLASGGAPLSKRTIGAVARFNLPCVROGYGILETTSAIIITPBCDDK	360
Db	301	ELLDKLYDLSNVLVJLASGGAPLSKRTIGAVARFNLPCVROGYGILETTSAIIITPBCDDK	360
Qy	361	PGASGKVVPLFKAKVIDLDTKKTETGPRRBEVCVCKGMLKKGVNDNEATREIIDEEGWL	420
Db	361	PGASGKVVPLFKAKVIDLDTKKTSGPRRBEVCVCKGMLKKGVNNNEATREIIDEEGWL	420
Qy	421	HTGDIGYVDEEKHEFIYDRLKSLIKYGYGVPPAELESVLLQHPNIFPDAGVAGVPPJAG	480
Db	421	HTGDIGYVDEEKHEFIYDRLKSLIKYGYGVPPAELESVLLQHPNIFPDAGVAGVPPVAG	480
Qy	481	ELPBAVVVYLKKGSMTEKEVMDYVASOVSAKRLRGVRFVDEVPKGLTGIDGKAIREI	540
Db	481	ELPBAVVVLESGKMTKEVMDYVASOVSAKRLRGVRFVDEVPKGLTGIDGKAIREI	540
Qy	541	LKKRVAKM 548	
Db	541	LKKRVAKM 548	

RESULT 15
US-09-111-752-7
Sequence 7, Application US/09111752-7
Patent No. 6074859
GENERAL INFORMATION:
APPLICANT: HIROKAWA, KOZO
APPLICANT: KATAYAMA, NAOKI
APPLICANT: MURAKAMI, SEIJI

```

: TITLE OF INVENTION: MUTANT-TYPE BIOLUMINESCENT PROTEIN, AND
: NUMBER OF INVENTION: PROCESS FOR PRODUCING MUTANT-TYPE BIOLUMINESCENT PROTEIN
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESS: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
: ADDRESS: P.C. JEFFERSON DAVIS HIGHWAY, SUITE 400
: STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
: CITY: ARLINGTON
: STATE: VA
: COUNTRY: USA
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/111,752
: FILING DATE: 08-JUL-1998
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: OBLON, NORMAN F.
: REGISTRATION NUMBER: 24,618
: REFERENCE/DOCKET NUMBER: 7126-0009-0
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-413-3000
: TELEFAX: 703-413-2220
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 552 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: ORIGINAL SOURCE:
: ORGANISM: Luciola cruciata and Photinus pyralis
: US-09-111-752-7

```

```

Query Match          91.7%; Score 2588; DB 3; Length 552;
Best Local Similarity 90.8%; Pred. No. 5.3e-262;
Matches 493; Conservative 32; Mismatches 18; Indels 0; Gaps 0;

```

```

QY 1 MNNENDNIIYGPPEPIPIEGSAGAOIIRKYMDRYAKGAIATNALTGVDYTAAYLE 60
   ||||| 11:|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1 MNNENDNIIYGPPEPIPIEGSAGAOIIRKYMDRYAKGAIATNALTGVDYTAAYLE 60
QY 61 KSCCIGALKNYGLVVDGRIALCSENCEEPFIVLAGIFGVGAPPTNETYTTLRELVHSL 120
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 61 KSCCIGALKNYGLVVDGRIALCSENCEEPFIVLAGIFGVGAPPTNETYTTLRELVHSL 120
QY 121 GISKPTIVFSSKGLDKVITVQKTIVTAIKTIVILDSKVYRGYSMDNFIKKNPQGFKG 180
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 121 GISKPTIVFSSKGLDKVITVQKTIVTAIKTIVILDSKVYRGYSMDNFIKKNPQGFKG 180
QY 181 SSFETVENRKREQVALLMNSSGSTGLPKGVQLTHEMLVTRSHARDPIYGNQVSPGTAIL 240
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 181 SSFETVENRKREQVALLMNSSGSTGLPKGVQLTHEMLVTRSHARDPIYGNQVSPGTAIL 240
QY 241 TVVPEHHGFGMTTLGLYTCGFRIVMLTKFDEETFLKTLDYKCSSVILVPTLFAILLNRS 300
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 241 TVVPEHHGFGMTTLGLYTCGFRIVMLTKFDEETFLKTLDYKCSSVILVPTLFAILLNRS 300
QY 301 ELLDKYDLSNLVETASGAPLSKEIGEAVARFNLPGVRQYGLTETTSAILITPEGDDK 360
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 301 ELLDKYDLSNLVETASGAPLSKEIGEAVARFNLPGVRQYGLTETTSAILITPEGDDK 360
QY 361 PGASGKVVPLFKAVIDLDDTKKTLGPNRGEVCYKGMILMKGYNDPNEATREITIDEGWL 420
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 361 PGASGKVVPLFKAVIDLDDTKKTLGPNRGEVCYKGMILMKGYNDPNEATREITIDEGWL 420
QY 421 HTGDIGYYDEEKHFFIVDRILKSLIKYGYOVPALESVLLQHPNIFDAGVAGLPDDAG 480
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

Db 421 HTGDIGYYDEEKHFFIVDRILKSLIKYGYOVPALESVLLQHPNIFDAGVAGLPDDAG 480
QY 481 ELPGAAYVYLKKGKSMTEKEVMDIVASOVSNAKRLRGVRFVDEVPKGLGKIDGKAIREI 540
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 481 ELPGAAYVYLKKGKSMTEKEVMDIVASOVSNAKRLRGVRFVDEVPKGLGKIDGKAIREI 540
QY 541 LKK 543
   |||
Db 541 LKK 543

```

```

Search completed: November 29, 2001, 01:53:47
Job time: 3207 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2001, 01:53:47 ; Search time 26.42 Seconds
(without alignments)
466.760 Million cell updates/sec

Title: US-09-581-241-6

Perfect score: 2823

Sequence: 1 MENMENDNIYGEPEFYPI.....TGKIDKAIKREILKKPVAKM 548

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued_patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	2817	99.8	548	2	US-08-460-934-2
2	2817	99.8	548	2	US-08-782-118-2
3	2815	99.7	548	3	US-09-111-752-14
4	2814	99.7	548	1	US-07-903-047-8
5	2814	99.7	548	4	US-09-380-061B-16
6	2812	99.6	636	2	US-08-460-934-9
7	2812	99.6	636	2	US-08-782-118-9
8	2799	99.1	568	2	US-08-460-934-6
9	2799	99.1	568	2	US-08-782-118-6
10	2701	95.7	552	3	US-09-111-752-10
11	2684	95.1	548	1	US-07-675-211-2
12	2684	95.1	548	1	US-07-903-047-2
13	2684	95.1	548	1	US-08-076-042-2
14	2684	95.1	548	1	US-09-380-061B-14
15	2590	91.7	552	3	US-09-111-752-7
16	2394	84.8	548	4	US-09-380-061B-18
17	2335.5	82.7	548	4	US-09-380-061B-20
18	1967.5	69.7	550	1	US-08-354-240A-4
19	1956.5	69.3	550	1	US-08-354-240A-6
20	1947.5	68.9	549	1	US-08-354-240A-2
21	1945.5	68.9	550	3	US-08-867-352-23
22	1945.5	68.9	550	4	US-09-380-061B-21
23	1945.5	68.9	815	1	US-08-123-520C-9
24	1942.5	68.8	550	4	US-08-718-425-5
25	1940.5	68.7	550	4	US-08-875-277A-4
26	1940.5	68.7	550	4	US-08-875-277A-2
27	1939.5	68.7	550	4	US-08-718-425-2

28	1939.5	68.7	550	4	US-08-875-277A-2	Sequence 2, Appl1
29	1937.5	68.6	550	4	US-09-380-061B-6	Sequence 6, Appl1
30	1905.5	67.5	561	2	US-08-474-169-8	Sequence 8, Appl1
31	1831	64.9	552	1	US-08-231-729B-6	Sequence 6, Appl1
32	723.5	25.6	540	4	US-08-991-677-8	Sequence 8, Appl1
33	587	20.8	578	4	US-08-981-215-1	Sequence 1, Appl1
34	260.5	9.2	798	2	US-08-222-617A-8	Sequence 20, Appl1
35	256	9.1	582	3	US-08-403-852D-20	Sequence 20, Appl1
36	256	9.1	582	3	US-08-510-646B-21	Sequence 20, Appl1
37	256	9.1	582	4	US-09-231-818-20	Sequence 2, Appl1
38	251.5	8.9	669	2	US-07-861-800-2	Sequence 2, Appl1
39	251	8.9	2618	4	US-09-413-814-28	Sequence 28, Appl1
40	249	8.8	2628	4	US-09-413-814-11	Sequence 11, Appl1
41	245	8.7	3665	2	US-08-222-617A-13	Sequence 13, Appl1
42	245	8.7	3712	2	US-08-222-617A-4	Sequence 4, Appl1
43	245	8.7	3712	2	US-08-222-617A-25	Sequence 25, Appl1
44	233	8.3	768	2	US-08-222-617A-5	Sequence 5, Appl1
45	233	8.3	3666	2	US-08-222-617A-12	Sequence 12, Appl1

ALIGNMENTS

RESULT 1
US-08-460-934-2
; Sequence 2, Application US/08460934
; Patent No. 5814465
; GENERAL INFORMATION:
; APPLICANT: TATSUMI, HIROKI
; APPLICANT: FUKUDA, SATOSHI
; APPLICANT: KIKUCHI, MAMORU
; APPLICANT: KOTAMA, YASUJI
; TITLE OF INVENTION: BIOTINATED FIREFLY LUCIFERASE, A GENE
; TITLE OF INVENTION: FOR BIOTINATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
; TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINATED AND A BIOLUMINESCENT
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,
; ADDRESS: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,934
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: A35
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 193798/1994
; FILING DATE: 27-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 54625/1995
; FILING DATE: 14-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 98857/1995
; FILING DATE: 24-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids

TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Luciola lateralis
 US-08-460-934-2

Query Match 99.8%; Score 2817; DB 2; Length 548;
 Best local similarity 99.6%; Pred. No. 6.6e-286;
 Matches 546; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MENMENDENIVGPEPEPYPIEESGAGOLRKRYAKIAGIAFTNALTVDTYAYYLE 60
 DB 1 MENMENDENIVGPEPEPYPIEESGAGOLRKRYAKIAGIAFTNALTVDTYAYYLE 60
 QY 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFFIPLVAGLFIGVAVAPNEIYTLRELHSL 120
 DB 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFFIPLVAGLFIGVAVAPNEIYTLRELHSL 120
 QY 121 GISKPTIVSSKKGDLKVTVOKTVAIKTIYILDSKVDYRGYOSMDNFKKNTPOGFGK 180
 DB 121 GISKPTIVSSKKGDLKVTVOKTVAIKTIYILDSKVDYRGYOSMDNFKKNTPOGFGK 180
 QY 181 SSFKTYEVNRKEOVALLIMSSGSTGLPKGVOLTHENIYTRFSHARDPIYGNVSPGTAIL 240
 DB 181 SSFKTYEVNRKEOVALLIMSSGSTGLPKGVOLTHENIYTRFSHARDPIYGNVSPGTAIL 240
 QY 241 TVPFFHGFMTTLGYLTCGFRIVMLTKFDEETFLKTLQDYKCSSVILVPTLFAILNRS 300
 DB 241 TVPFFHGFMTTLGYLTCGFRIVMLTKFDEETFLKTLQDYKCSSVILVPTLFAILNRS 300
 QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRNPLPGVROGGLTETTSAILITPEGDDK 360
 DB 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRNPLPGVROGGLTETTSAILITPEGDDK 360
 QY 361 ELDDKYDLSNLVEIASGAPLSKEIGEAVARRNPLPGVROGGLTETTSAILITPEGDDK 420
 DB 361 ELDDKYDLSNLVEIASGAPLSKEIGEAVARRNPLPGVROGGLTETTSAILITPEGDDK 420
 QY 361 PGASGVVPLFAKAYIDLTDTKTLGPNRRGEVCGVPMKGYVNDPEATREIIDEEGWL 480
 DB 361 PGASGVVPLFAKAYIDLTDTKTLGPNRRGEVCGVPMKGYVNDPEATREIIDEEGWL 480
 QY 421 HNGDIGYDEKHEFTVDRLSLTKRYGYOVPALESVLLQHPNLEFDGAVAPDPPIAG 480
 DB 421 HNGDIGYDEKHEFTVDRLSLTKRYGYOVPALESVLLQHPNLEFDGAVAPDPPIAG 480
 QY 481 ELPGAVVYLKKGKSMTEKEMVDYASOVNAKRLRGVRFVDEVPKGLTGIDGKAIREI 540
 DB 481 ELPGAVVYLKKGKSMTEKEMVDYASOVNAKRLRGVRFVDEVPKGLTGIDGKAIREI 540
 QY 541 LKKPVAKM 548
 DB 541 LKKPVAKM 548

RESULT 2
 US-08-782-118-2
 Sequence 2, Application US/08782118
 Patent No. 5843746

GENERAL INFORMATION:
 APPLICANT: TATSUMI, HIROKI
 APPLICANT: FUKUDA, SATOSHI
 APPLICANT: KIKUCHI, MAMORU
 APPLICANT: KOYAMA, YASUJI
 TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
 TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
 TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
 ADDRESSEE: P. C. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
 STREET: 175 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
 CITY: ARLINGTON

STATE: VA
 COUNTRY: USA
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/782,118
 FILING DATE: 13-JAN-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/460,934
 FILING DATE: 05-JUN-1995
 APPLICATION NUMBER: JP 193798/1994
 FILING DATE: 27-JUL-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 54625/1995
 FILING DATE: 14-MAR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 98857/1995
 FILING DATE: 24-APR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: OBLON, NORMAN F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 7126-001-0
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 548 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Luciola lateralis
 US-08-782-118-2

Query Match 99.8%; Score 2817; DB 2; Length 548;
 Best local similarity 99.6%; Pred. No. 6.6e-286;
 Matches 546; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MENMENDENIVGPEPEPYPIEESGAGOLRKRYAKIAGIAFTNALTVDTYAYYLE 60
 DB 1 MENMENDENIVGPEPEPYPIEESGAGOLRKRYAKIAGIAFTNALTVDTYAYYLE 60
 QY 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFFIPLVAGLFIGVAVAPNEIYTLRELHSL 120
 DB 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFFIPLVAGLFIGVAVAPNEIYTLRELHSL 120
 QY 121 GISKPTIVSSKKGDLKVTVOKTVAIKTIYILDSKVDYRGYOSMDNFKKNTPOGFGK 180
 DB 121 GISKPTIVSSKKGDLKVTVOKTVAIKTIYILDSKVDYRGYOSMDNFKKNTPOGFGK 180
 QY 181 SSFKTYEVNRKEOVALLIMSSGSTGLPKGVOLTHENIYTRFSHARDPIYGNVSPGTAIL 240
 DB 181 SSFKTYEVNRKEOVALLIMSSGSTGLPKGVOLTHENIYTRFSHARDPIYGNVSPGTAIL 240
 QY 241 TVPFFHGFMTTLGYLTCGFRIVMLTKFDEETFLKTLQDYKCSSVILVPTLFAILNRS 300
 DB 241 TVPFFHGFMTTLGYLTCGFRIVMLTKFDEETFLKTLQDYKCSSVILVPTLFAILNRS 300
 QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRNPLPGVROGGLTETTSAILITPEGDDK 360
 DB 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRNPLPGVROGGLTETTSAILITPEGDDK 360
 QY 361 PGASGVVPLFAKAYIDLTDTKTLGPNRRGEVCGVPMKGYVNDPEATREIIDEEGWL 420
 DB 361 PGASGVVPLFAKAYIDLTDTKTLGPNRRGEVCGVPMKGYVNDPEATREIIDEEGWL 420

QY 421 HTGDIGYDEKHFIVDRLSLITKYGQVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
 DB 421 HTGDIGYDEKHFIVDRLSLITKYGQVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
 QY 481 ELPGAVVVLKKGSMTEKEVMDVYASQVSNARKRLRGVRFVDEVPKGLTGKIDGKAIREI 540
 DB 481 ELPGAVVVLKKGSMTEKEVMDVYASQVSNARKRLRGVRFVDEVPKGLTGKIDGKAIREI 540
 QY 541 LKRPVAKM 548
 DB 541 LKRPVAKM 548
 RESULT 3
 US-09-111-752-14
 ; Sequence 14, Application US/09111752
 ; Patent No. 6074859
 ; GENERAL INFORMATION:
 ; APPLICANT: HIROKAWA, KOZO
 ; APPLICANT: KAJIYAMA, NAOKI
 ; APPLICANT: MURAKAMI, SEIJI
 ; TITLE OF INVENTION: MUTANT-TYPE BIOLUMINESCENT PROTEIN, AND
 ; TITLE OF INVENTION: PROCESS FOR PRODUCING MUTANT-TYPE LUMINESCENT PROTEIN
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 ; ADDRESSEE: P.C.
 ; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
 ; CITY: ARLINGTON
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22202
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/111,752
 ; FILING DATE: 08-JUL-1998
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: OBLON, NORMAN F.
 ; REGISTRATION NUMBER: 24,618
 ; REFERENCE/DOCKET NUMBER: 7126-0009-0
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 703-413-3000
 ; TELEFAX: 703-413-2220
 ; INFORMATION FOR SEQ ID NO: 14:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 548 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-09-111-752-14
 Query Match 99.7%; Score 2815; DB 3; Length 548;
 Best Local Similarity 99.6%; Pred. No. 1.1e-285;
 Matches 546; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

DB 121 GISKPTIVSSKGLDKVTVQKTVTAIKTIVILDSKVDRGYSMDNFIRKNTPOGFRG 180
 QY 181 SSFKTVEVNRKQVALLINSSGSTGLPKGVOLTHENITVTRSHADPIYGNQVSGTAIL 240
 DB 181 SSFKTVEVNRKQVALLINSSGSTGLPKGVOLTHENITVTRSHADPIYGNQVSGTAIL 240
 QY 241 TVVPRFHGFMFTTGLTGLCGFRIMVLTFRDEETFLKTLQDYKCSSVILVPLFAILNRS 300
 DB 241 TVVPRFHGFMFTTGLTGLCGFRIMVLTFRDEETFLKTLQDYKCSSVILVPLFAILNRS 300
 QY 301 ELIDKXDSLNLVEIASGAPLSKEIGEAARRRNDPGVQGLTETTSAILITPEGDGK 360
 DB 301 ELIDKXDSLNLVEIASGAPLSKEIGEAARRRNDPGVQGLTETTSAILITPEGDGK 360
 QY 361 PGASGVVPLFPAKVIDLDTKTLGPNRGEVCGVPMILMKGYVNPETRETIIDEEGML 420
 DB 361 PGASGVVPLFPAKVIDLDTKTLGPNRGEVCGVPMILMKGYVNPETRETIIDEEGML 420
 QY 421 HTGDIGYDEKHFIVDRLSLITKYGQVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
 DB 421 HTGDIGYDEKHFIVDRLSLITKYGQVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
 QY 481 ELPGAVVVLKKGSMTEKEVMDVYASQVSNARKRLRGVRFVDEVPKGLTGKIDGKAIREI 540
 DB 481 ELPGAVVVLKKGSMTEKEVMDVYASQVSNARKRLRGVRFVDEVPKGLTGKIDGKAIREI 540
 QY 541 LKRPVAKM 548
 DB 541 LKRPVAKM 548
 RESULT 4
 US-07-903-047-8
 ; Sequence 8, Application US/07903047
 ; Patent No. 5229285
 ; GENERAL INFORMATION:
 ; APPLICANT: Kajiya, Naoki
 ; APPLICANT: Nakano, Eiichi
 ; TITLE OF INVENTION: Thermostable Luciferase of Firefly,
 ; TITLE OF INVENTION: DNA, And Process For The Preparation Of Thermostable
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/903,047
 ; FILING DATE: 19920623
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mastrock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 7005-048
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 790-9090
 ; TELEFAX: 212 869-8864/9741
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 548 amino acids
 ; TYPE: AMINO ACID
 ; STRANDEDNESS: single
 ; TOPOLOGY: unknown

MOLECULE TYPE: Peptide
US-07-903-047-8

Query Match 99.7%; Score 2814; DB 1; Length 548;
Best Local Similarity 99.6%; Pred. No. 1.4e-285;
Matches 546; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MENMENDENIVYGEPEFYPIEESGAGAQLRKYMDRYAKGAIATFTNALGVDTYAYEYL 60
DB 1 MENMENDENIVYGEPEFYPIEESGAGAQLRKYMDRYAKGAIATFTNALGVDTYAYEYL 60
QY 61 KSCCCEALKNGVLVVDGRIALCSENCEEFPIVLAGLFTGVGAPVPTNFIYTLRELIVHSL 120
DB 61 KSCCCEALKNGVLVVDGRIALCSENCEEFPIVLAGLFTGVGAPVPTNFIYTLRELIVHSL 120
QY 121 GISKPTIVSSKKGLDVTYOKTVAIKTIYILDSKVYRGYSMDNFITKNTPOGFGK 180
DB 121 GISKPTIVSSKKGLDVTYOKTVAIKTIYILDSKVYRGYSMDNFITKNTPOGFGK 180
QY 181 SSFKTYEVNRRKQVALINMSSGSTGLPKGVOLTHENIYTRFESHARDPIYGNVSPGTAIL 240
DB 181 SSFKTYEVNRRKQVALINMSSGSTGLPKGVOLTHENIYTRFESHARDPIYGNVSPGTAIL 240
QY 241 TVVPRHHGFMFTTGLYLCGFRIVMLTKFDEETFLKTLQDYKCSVILVPTLFAILNRS 300
DB 241 TVVPRHHGFMFTTGLYLCGFRIVMLTKFDEETFLKTLQDYKCSVILVPTLFAILNRS 300
QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRPNLPGVROGVLTEETSAIITPEGDDK 360
DB 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRPNLPGVROGVLTEETSAIITPEGDDK 360
QY 361 PGASGVVPLFRAKYIDDTKTKTLGNRRGEVCGVPMIMKGVDPNPEATREIIDEEGML 420
DB 361 PGASGVVPLFRAKYIDDTKTKTLGNRRGEVCGVPMIMKGVDPNPEATREIIDEEGML 420
QY 421 HTGDIGYDEEKNFTIVDRLSLIKYGQVPALEESVLLQHPNIFDAGVAGVDPPIAG 480
DB 421 HTGDIGYDEEKNFTIVDRLSLIKYGQVPALEESVLLQHPNIFDAGVAGVDPPIAG 480
QY 481 ELPGAIVVLLKKGKSMTEKEVMDYVASOVSNARKLRGVRVDEVKGLTGKIDGKAIREI 540
DB 481 ELPGAIVVLLKKGKSMTEKEVMDYVASOVSNARKLRGVRVDEVKGLTGKIDGKAIREI 540
QY 541 LKRPVAKM 548
DB 541 LKRPVAKM 548

```

RESULT 5
US-09-380-061B-16
Sequence 16, Application US/09380061B
Patent No. 6265177

GENERAL INFORMATION:

APPLICANT: SOUTIRRELL, DAVID JAMES
WHITE, PETER JOHN
LOWE, CHRISTOPHER ROBIN
MURRAY, JAMES AUGUSTUS HENRY
TITLE OF INVENTION: ENZYME ASSAY FOR MUTANT FIREFLY LUCIFERASE
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/380,061B
FILING DATE: 25-Aug-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB98/01026
FILING DATE: 7-APR-1998
APPLICATION NUMBER: GB 9707468.8
FILING DATE: 11-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B. J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 124-725
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4000
TELEFAX: (703)816-4100
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-380-061B-16

Query Match 99.7%; Score 2814; DB 4; Length 548;
Best Local Similarity 99.6%; Pred. No. 1.4e-285;
Matches 546; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MENMENDENIVYGEPEFYPIEESGAGAQLRKYMDRYAKGAIATFTNALGVDTYAYEYL 60
DB 1 MENMENDENIVYGEPEFYPIEESGAGAQLRKYMDRYAKGAIATFTNALGVDTYAYEYL 60
QY 61 KSCCCEALKNGVLVVDGRIALCSENCEEFPIVLAGLFTGVGAPVPTNFIYTLRELIVHSL 120
DB 61 KSCCCEALKNGVLVVDGRIALCSENCEEFPIVLAGLFTGVGAPVPTNFIYTLRELIVHSL 120
QY 121 GISKPTIVSSKKGLDVTYOKTVAIKTIYILDSKVYRGYSMDNFITKNTPOGFGK 180
DB 121 GISKPTIVSSKKGLDVTYOKTVAIKTIYILDSKVYRGYSMDNFITKNTPOGFGK 180
QY 181 SSFKTYEVNRRKQVALINMSSGSTGLPKGVOLTHENIYTRFESHARDPIYGNVSPGTAIL 240
DB 181 SSFKTYEVNRRKQVALINMSSGSTGLPKGVOLTHENIYTRFESHARDPIYGNVSPGTAIL 240
QY 241 TVVPRHHGFMFTTGLYLCGFRIVMLTKFDEETFLKTLQDYKCSVILVPTLFAILNRS 300
DB 241 TVVPRHHGFMFTTGLYLCGFRIVMLTKFDEETFLKTLQDYKCSVILVPTLFAILNRS 300
QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRPNLPGVROGVLTEETSAIITPEGDDK 360
DB 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRPNLPGVROGVLTEETSAIITPEGDDK 360
QY 361 PGASGVVPLFRAKYIDDTKTKTLGNRRGEVCGVPMIMKGVDPNPEATREIIDEEGML 420
DB 361 PGASGVVPLFRAKYIDDTKTKTLGNRRGEVCGVPMIMKGVDPNPEATREIIDEEGML 420
QY 421 HTGDIGYDEEKNFTIVDRLSLIKYGQVPALEESVLLQHPNIFDAGVAGVDPPIAG 480
DB 421 HTGDIGYDEEKNFTIVDRLSLIKYGQVPALEESVLLQHPNIFDAGVAGVDPPIAG 480
QY 481 ELPGAIVVLLKKGKSMTEKEVMDYVASOVSNARKLRGVRVDEVKGLTGKIDGKAIREI 540
DB 481 ELPGAIVVLLKKGKSMTEKEVMDYVASOVSNARKLRGVRVDEVKGLTGKIDGKAIREI 540
QY 541 LKRPVAKM 548
DB 541 LKRPVAKM 548

```

RESULT 6
US-08-460-934-9
Sequence 9, Application US/08460934

```

Patent No. 5814465
GENERAL INFORMATION:
APPLICANT: TATSUMI, HIROKI
APPLICANT: FUKUDA, SATOSHI
APPLICANT: KIKUCHI, MAMORU
APPLICANT: Koyama, Yasuji
TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT
TITLE OF INVENTION: ANALYSIS METHOD
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,934
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 193798/1994
FILING DATE: 27-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 54625/1995
FILING DATE: 14-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 98857/1995
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 636 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-934-9

```

```

Query Match          99.6%; Score 2812; DB 2; Length 636;
Best Local Similarity 99.6%; Pred. No. 2.9e-285;
Matches 545; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MENMENDENIVYGPPEFYPIEGSGAGQIRKYMDRKAGIAFTNALTGVDYTAETLE 60
DB 1 MENMENDENIVYGPPEFYPIEGSGAGQIRKYMDRKAGIAFTNALTGVDYTAETLE 60
QY 61 KSCCLGEALKNYGLVVDGRIALCSECEFFIPVLAGLFTIGVVAFTNEIYTLRELHVS 120
DB 61 KSCCLGEALKNYGLVVDGRIALCSECEFFIPVLAGLFTIGVVAFTNEIYTLRELHVS 120
QY 121 GISKPTIVSSKKGDKLVTVQKTYATITVILDSKVDYRGYQSDNFKTKNTQGFPG 180
DB 121 GISKPTIVSSKKGDKLVTVQKTYATITVILDSKVDYRGYQSDNFKTKNTQGFPG 180
QY 181 SSKPTVEVRKREOVALLIMSSSGTGPKGVOLTHENITVRFSHARDPIYGNQVSPGTA 240
DB 181 SSKPTVEVRKREOVALLIMSSSGTGPKGVOLTHENITVRFSHARDPIYGNQVSPGTA 240

```

```

QY 241 TVPPEHHGFMFTFLGILYLCGFRIYMLTKFDEETFLKTLQDYKCSSVILVPTLAILNRS 300
DB 241 TVPPEHHGFMFTFLGILYLCGFRIYMLTKFDEETFLKTLQDYKCSSVILVPTLAILNRS 300
QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGYRQGYGLTETTSAILITPEGDK 360
DB 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGYRQGYGLTETTSAILITPEGDK 360
QY 361 PGASGVVPLFAKAVYDLDTKTKLGPNRGECYVGPMLMKGYVNPATREIIDEEGML 420
DB 361 PGASGVVPLFAKAVYDLDTKTKLGPNRGECYVGPMLMKGYVNPATREIIDEEGML 420
QY 421 HTGIDIGYDEKHFLLVDRLKSLIKKGYVPPAELESYLLQHPNIFDAGVAGVDPPIAG 480
DB 421 HTGIDIGYDEKHFLLVDRLKSLIKKGYVPPAELESYLLQHPNIFDAGVAGVDPPIAG 480
QY 481 ELPGAVVYLKKGKSMTEKEVMYVASQVSNARKRGVRFVDEVPKGLTKIDGKAIREI 540
DB 481 ELPGAVVYLKKGKSMTEKEVMYVASQVSNARKRGVRFVDEVPKGLTKIDGKAIREI 540
QY 541 LKKPVAK 547
DB 541 LKKPVAK 547

```

```

RESULT 7
US-08-782-118-9
Sequence 9, Application US/08782118
Patent No. 5843746
GENERAL INFORMATION:
APPLICANT: TATSUMI, HIROKI
APPLICANT: FUKUDA, SATOSHI
APPLICANT: KIKUCHI, MAMORU
APPLICANT: Koyama, Yasuji
TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT
TITLE OF INVENTION: ANALYSIS METHOD
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/782,118
FILING DATE: 13-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/460,934
FILING DATE: 05-JUN-1995
APPLICATION NUMBER: JP 193798/1994
FILING DATE: 27-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 54625/1995
FILING DATE: 14-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 98857/1995
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000

```

TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 636 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-782-118-9

Query Match 99.6%; Score 2812; DB 2; Length 636;
 Best Local Similarity 99.6%; Pred. No. 2.9e-285;
 Matches 545; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MENMENDENIVGPEPFYPIEESAGAOQLRKYMDRAKIGALFTVALGVDYTAEYLE 60
DB 1 MENMENDENIVGPEPFYPIEESAGAOQLRKYMDRAKIGALFTVALGVDYTAEYLE 60
QY 61 KSCCLGEBALKNGLVVDGRALALSCENCEEFPIVLAGLFGVAPPTNEITLRELVHSL 120
DB 61 KSCCLGEBALKNGLVVDGRALALSCENCEEFPIVLAGLFGVAPPTNEITLRELVHSL 120
QY 121 GISKPTIVSSKKGDLKVTIVOKTVAITIVITLDSKVYRGVQSMDFIKKNTPOGFGK 180
DB 121 GISKPTIVSSKKGDLKVTIVOKTVAITIVITLDSKVYRGVQSMDFIKKNTPOGFGK 180
QY 181 SSKFTVEVNRKEQVALINMSSGSTGLPKGVOLTHENITVRFSHARDPIYGNVSPGTAIL 240
DB 181 SSKFTVEVNRKEQVALINMSSGSTGLPKGVOLTHENITVRFSHARDPIYGNVSPGTAIL 240
QY 241 TVVPHFGFMFTTIGYLTGCGRIYMLTFDEDEFLKTIQDYKCSVILVPLFAILNRS 300
DB 241 TVVPHFGFMFTTIGYLTGCGRIYMLTFDEDEFLKTIQDYKCSVILVPLFAILNRS 300
QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGSGLTETTSAILITPEGDOK 360
DB 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGSGLTETTSAILITPEGDOK 360
QY 361 PGASGVVPLFAKVIDLDTKTLGPNRGEVCGKPMIMKGYVNDPEATREIIDEESGWL 420
DB 361 PGASGVVPLFAKVIDLDTKTLGPNRGEVCGKPMIMKGYVNDPEATREIIDEESGWL 420
QY 421 HNGDIGYDEEKNFTVDRKLSLIKYGQVPPALESVLLQHPNIFPAGVAGVDDPTAG 480
DB 421 HNGDIGYDEEKNFTVDRKLSLIKYGQVPPALESVLLQHPNIFPAGVAGVDDPTAG 480
QY 481 ELPGAVVVLKKGKSMTEKEVMDYVASQVSNARKLGGVRFVDEVPKGLTGKIDKAIKEI 540
DB 481 ELPGAVVVLKKGKSMTEKEVMDYVASQVSNARKLGGVRFVDEVPKGLTGKIDKAIKEI 540
QY 541 LKKPYAK 547
DB 541 LKKPYAK 547

```

RESULT 8
 US-08-460-934-6
 Sequence 6, Application US/08460934
 Patent No. 5814465

GENERAL INFORMATION:
 APPLICANT: TATSUMI, HIROKI
 APPLICANT: FUKUDA, SATOSHI
 APPLICANT: KIRUCHI, MAMORU
 APPLICANT: KOYAMA, YASUJI
 TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
 TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
 TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 ADDRESSEE: P.C.
 STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR

CITY: ARLINGTON
 STATE: VA
 COUNTRY: USA
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patentln Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/460,934
 FILING DATE: 05-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 193798/1994
 FILING DATE: 27-JUL-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 54625/1995
 FILING DATE: 14-MAR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 98857/1995
 FILING DATE: 24-APR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: OBLON, NORMAN F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 7126-001-0
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 568 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-460-934-6

Query Match 99.1%; Score 2799; DB 2; Length 568;
 Best Local Similarity 99.3%; Pred. No. 5.4e-284;
 Matches 542; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

QY 3 MENMENDENIVGPEPFYPIEESAGAOQLRKYMDRAKIGALFTVALGVDYTAEYLE 62
DB 3 MENMENDENIVGPEPFYPIEESAGAOQLRKYMDRAKIGALFTVALGVDYTAEYLE 62
QY 63 CCLGEBALKNGLVVDGRALALSCENCEEFPIVLAGLFGVAPPTNEITLRELVHSLGI 122
DB 63 CCLGEBALKNGLVVDGRALALSCENCEEFPIVLAGLFGVAPPTNEITLRELVHSLGI 122
QY 123 SKPTIVSSKKGDLKVTIVOKTVAITIVITLDSKVYRGVQSMDFIKKNTPOGFGKSS 182
DB 123 SKPTIVSSKKGDLKVTIVOKTVAITIVITLDSKVYRGVQSMDFIKKNTPOGFGKSS 182
QY 183 SKPTIVSSKKGDLKVTIVOKTVAITIVITLDSKVYRGVQSMDFIKKNTPOGFGKSS 202
DB 183 SKPTIVSSKKGDLKVTIVOKTVAITIVITLDSKVYRGVQSMDFIKKNTPOGFGKSS 202
QY 203 FKTVEVNRKEQVALINMSSGSTGLPKGVOLTHENITVRFSHARDPIYGNVSPGTAILTV 262
DB 203 FKTVEVNRKEQVALINMSSGSTGLPKGVOLTHENITVRFSHARDPIYGNVSPGTAILTV 262
QY 243 VPFHFGFMFTTIGYLTGCGRIYMLTFDEDEFLKTIQDYKCSVILVPLFAILNRSGL 302
DB 243 VPFHFGFMFTTIGYLTGCGRIYMLTFDEDEFLKTIQDYKCSVILVPLFAILNRSGL 302
QY 303 LDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGSGLTETTSAILITPEGDDKPG 362
DB 303 LDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGSGLTETTSAILITPEGDDKPG 362
QY 363 ASKVPLFAKVIDLDTKTLGPNRGEVCGKPMIMKGYVNDPEATREIIDEESGWLHT 422
DB 363 ASKVPLFAKVIDLDTKTLGPNRGEVCGKPMIMKGYVNDPEATREIIDEESGWLHT 422
QY 423 GDIGYDEEKNFTVDRKLSLIKYGQVPPALESVLLQHPNIFPAGVAGVDDPTAGEL 482
DB 423 GDIGYDEEKNFTVDRKLSLIKYGQVPPALESVLLQHPNIFPAGVAGVDDPTAGEL 482
QY 443 GDIGYDEEKNFTVDRKLSLIKYGQVPPALESVLLQHPNIFPAGVAGVDDPTAGEL 502
DB 443 GDIGYDEEKNFTVDRKLSLIKYGQVPPALESVLLQHPNIFPAGVAGVDDPTAGEL 502

```

	Qy	Db
483	PGAVVVLKKGKSMTEKEMVDVASQVSNARKLRGVRFVDEVPGLGKIDGKAIRELK	54.2
503	PGAVVVLKKGKSMTEKEMVDVASQVSNARKLRGVRFVDEVPGLGKIDGKAIRELK	56.2
543	KPVAKM	54.8
563	KPVAKM	56.8

RESULT 9
US-08-782-118-6
; Sequence 6, Application US/08782118
; Patent No. 5843746

APPLICANT: TATSUMI, HIROKI
 APPLICANT: FUKUDA, SATOSHI
 APPLICANT: KIKUCHI, MAMORU
 APPLICANT: Koyama, YASUJI
 TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
 TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA,
 TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINATED AND A BIOLUMINESCENT
 TITLE OF INVENTION: ANALYSIS METHOD
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 ADDRESSEE: P.C.
 STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
 CITY: ARLINGTON
 STATE: VA
 COUNTRY: USA
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/782,118
 FILING DATE: 13-JAN-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/460,934
 FILING DATE: 05-JUN-1995
 APPLICATION NUMBER: JP 193798/1994
 FILING DATE: 27-JUL-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 54625/1995
 FILING DATE: 14-MAR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 98857/1995
 FILING DATE: 24-APR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: OBLON, NORMAN F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 7126-001-0
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 568 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-782-118-6

Query Match	99.1%;	Score 2799;	DB 2;	Length 568;
Best Local Similarity	99.3%;	Pred. No. 5,4e-204;		
Matches 542;	Conservative	4;	Mismatches	0;
			Indels	0;
			Gaps	0;
3	NMEDENIYVGPFPPIEGSSAGQQLRKYMDRYAKGLAIFATNALGVDDTYAAEYLERS 62			

Db	23	SLDENDEIVIVGPPRPYDIEEGSAGADLRKMDRYATLGAIAFTNALTGVDYTYAEYLEKS	82
QY	63	CCJGEALKNGLVVDGRIALCSENCDEEFFIPVLAGLFIGVAVPMEITYTLRLVHSLGI	122
Db	83	CCJGEALKNGLVVDGRIALCSENCDEEFFIPVLAGLFIGVAVPMEITYTLRLVHSLGI	142
QY	123	SKPIPIVSSKKGDKVTYQKTYATLKITIYLDSKVDYRGQSMDFIKKNPQGRKSS	182
Db	143	SKPIPIVSSKKGDKVTYQKTYATLKITIYLDSKVDYRGQSMDFIKKNPQGRKSS	202
QY	183	FKTYEVARKEQVALINMSSGSTGLPKGVOLTHENITYRFSHARDPIYGNQVSGTALTY	242
Db	203	FKTYEVARKEQVALINMSSGSTGLPKGVOLTHENITYRFSHARDPIYGNQVSGTALTY	262
QY	243	VPFIHNGGMLFTTLGTYLNCGRPIYVMTLTFEDETFPLKTLQDYKCSSVILVPLFLINLSEL	302
Db	263	VPFIHNGGMLFTTLGTYLNCGRPIYVMTLTFEDETFPLKTLQDYKCSSVILVPLFLINLSEL	322
QY	303	LDKYDLSMLVEIASGAPALSKKEIGEAVARRENLPGVBOGYGLTETTSAILITPBGDDKPG	362
Db	323	LDKYDLSMLVEIASGAPALSKKEIGEAVARRENLPGVBOGYGLTETTSAILITPBGDDKPG	382
QY	363	ASGVAVPLFKAKVIDLDTKTKTLGPNRSGEYVCKPMLMKGVQVNDPEATRELIDEEGMLHT	422
Db	383	ASGVAVPLFKAKVIDLDTKTKTLGPNRSGEYVCKPMLMKGVQVNDPEATRELIDEEGMLHT	442
QY	423	GDIGYDEEKEHPEIYDRLSKLITKKGQVYPAELESYLLQHPNITFDGAVAGVDPPIAGEL	482
Db	443	GDIGYDEEKEHPEIYDRLSKLITKKGQVYPAELESYLLQHPNITFDGAVAGVDPPIAGEL	502
QY	483	PGAVVVLKKGKSMTEKEYEMDYVASOVSAKRLRGVAVDEVPRGGLTKIDGKAIARELIK	542
Db	503	PGAVVVLKKGKSMTEKEYEMDYVASOVSAKRLRGVAVDEVPRGGLTKIDGKAIARELIK	562
QY	543	KPVAKM	548
Db	563	KPVAKM	568

RESULT 10
 US-09-111-752-10
 ; Sequence 10, Application US/09111752
 ; Patent No. 6074859
 ; GENERAL INFORMATION:
 ; APPLICANT: HIROKAWA, KOZO
 ; APPLICANT: KAJIYAMA, NAKOI
 ; APPLICANT: MORAKAMI, SEIJI
 ; TITLE OF INVENTION: MUTANT-TYPE BIOLUMINESCENT PROTEIN, AND
 ; TITLE OF INVENTION: PROCESS FOR PRODUCING MUTANT-TYPE LUMINESCENT PROTEIN
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 ; ADDRESSEE: P.C.
 ; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
 ; CITY: ARLINGTON
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22202
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/111,752
 ; FILING DATE: 08-JUL-1998
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: OBLON, NORMAN F.
 ; REGISTRATION NUMBER: 24,618
 ; REFERENCE/DOCKET NUMBER: 7126-0009-0
 ; TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 552 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Luciola lateralis, Photinus pyralis
 US-09-111-752-10

Query Match 95.7%; Score 2701; DB 3; Length 552;
 Best Local Similarity 96.1%; Pred. No. 9.3e-274;
 Matches 522; Conservative 11; Mismatches 10; Indels 0; Gaps 0;

1 MENMENDENIVYGPPEPYPIEBSAGAOQLRKYMDRAKALGALFTNALGVDTYAEYLE 60
 1 MEMMENDENIVYGPPEPYPIEBSAGAOQLRKYMDRAKALGALFTNALGVDTYAEYLE 60
 61 KSCCLGALKNYGLVYDGRALCSENCCEFFIPVLAGLFIGVAPATNEIYTLRELHSL 120
 61 KSCCLGALKNYGLVYDGRALCSENCCEFFIPVLAGLFIGVAPATNEIYTLRELHSL 120
 121 GISKPTIVSSKKGLDKVTVTKYTAIKTIYILDSKVYRGYSMDNFIKKNTPQGFKG 180
 121 GISKPTIVSSKKGLDKVTVTKYTAIKTIYILDSKVYRGYSMDNFIKKNTPQGFKG 180
 181 SSFKTEVNRKKEQVALINSSSGTGLPKGVOLTHENIVTRFSHARDPIYGNVSPGTA 240
 181 SSFKTEVNRKKEQVALINSSSGTGLPKGVOLTHENIVTRFSHARDPIYGNVSPGTA 240
 241 TVVPHHGFMTTLLGYLTCGFRIVMLTKFDEETFLKTLQDYKCSSVILVPLFAILNRS 300
 241 TVVPHHGFMTTLLGYLTCGFRIVMLTKFDEETFLKTLQDYKCSSVILVPLFAILNRS 300
 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRENLPVROGYGLTETTSAILITPEGDDK 360
 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRENLPVROGYGLTETTSAILITPEGDDK 360
 361 PGASGVVPLFRAKVIDLDTKKTLPNRRGEVCVGPMLMGKYVDPNPAATREIIDEESGWL 420
 361 PGASGVVPLFRAKVIDLDTKKTLPNRRGEVCVGPMLMGKYVDPNPAATREIIDEESGWL 420
 421 HTGDIGYDEEKHFIVLRKSLIKYKGYVPALLESVLLQHPNIDAGVAGVDDPTAG 480
 421 HTGDIGYDEEKHFIVLRKSLIKYKGYVPALLESVLLQHPNIDAGVAGVDDPTAG 480
 481 ELPGAIVVLKKGKSKTEKEVMDYVAQVSNKRLRGVRFVDEVPKGLTGKIDKAIKREI 540
 481 ELPGAIVVLKKGKSKTEKEVMDYVAQVSNKRLRGVRFVDEVPKGLTGKIDKAIKREI 540
 541 LKK 543
 541 LKK 543

RESULT 11
 US-07-675-211-2
 Sequence 2, Application US/07675211
 Patent No. 5219737
 GENERAL INFORMATION:
 APPLICANT: KAJIYAMA, NAOKI
 APPLICANT: KAJIYAMA, EITICHI
 TITLE OF INVENTION: MUTANT LUCIFERASE OF A FIREFLY, MUTANT
 TITLE OF INVENTION: LUCIFERASE GENES, NOVEL RECOMBINANT DNAs CONTAINING THE
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: PENNIE & EDMONDS
 STREET: 1155 AVENUE OF THE AMERICAS

CITY: NEW YORK
 STATE: N.Y.
 COUNTRY: U.S.A.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/675,211
 FILING DATE: 19910326
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MISROCK, S. LESLIE
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 7005-026-999
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-790-9090
 TELEFAX: 212-869-9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 548 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Luciola cruciata
 US-07-675-211-2

Query Match 95.1%; Score 2684; DB 1; Length 548;
 Best Local Similarity 93.4%; Pred. No. 5.5e-272;
 Matches 512; Conservative 26; Mismatches 10; Indels 0; Gaps 0;

1 MENMENDENIVYGPPEPYPIEBSAGAOQLRKYMDRAKALGALFTNALGVDTYAEYLE 60
 1 MEMMENDENIVYGPPEPYPIEBSAGAOQLRKYMDRAKALGALFTNALGVDTYAEYLE 60
 61 KSCCLGALKNYGLVYDGRALCSENCCEFFIPVLAGLFIGVAPATNEIYTLRELHSL 120
 61 KSCCLGALKNYGLVYDGRALCSENCCEFFIPVLAGLFIGVAPATNEIYTLRELHSL 120
 121 GISKPTIVSSKKGLDKVTVTKYTAIKTIYILDSKVYRGYSMDNFIKKNTPQGFKG 180
 121 GISKPTIVSSKKGLDKVTVTKYTAIKTIYILDSKVYRGYSMDNFIKKNTPQGFKG 180
 181 SSFKTEVNRKKEQVALINSSSGTGLPKGVOLTHENIVTRFSHARDPIYGNVSPGTA 240
 181 SSFKTEVNRKKEQVALINSSSGTGLPKGVOLTHENIVTRFSHARDPIYGNVSPGTA 240
 241 TVVPHHGFMTTLLGYLTCGFRIVMLTKFDEETFLKTLQDYKCSSVILVPLFAILNRS 300
 241 TVVPHHGFMTTLLGYLTCGFRIVMLTKFDEETFLKTLQDYKCSSVILVPLFAILNRS 300
 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRENLPVROGYGLTETTSAILITPEGDDK 360
 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRENLPVROGYGLTETTSAILITPEGDDK 360
 361 PGASGVVPLFRAKVIDLDTKKTLPNRRGEVCVGPMLMGKYVDPNPAATREIIDEESGWL 420
 361 PGASGVVPLFRAKVIDLDTKKTLPNRRGEVCVGPMLMGKYVDPNPAATREIIDEESGWL 420
 421 HTGDIGYDEEKHFIVLRKSLIKYKGYVPALLESVLLQHPNIDAGVAGVDDPTAG 480
 421 HTGDIGYDEEKHFIVLRKSLIKYKGYVPALLESVLLQHPNIDAGVAGVDDPTAG 480
 481 ELPGAIVVLKKGKSKTEKEVMDYVAQVSNKRLRGVRFVDEVPKGLTGKIDKAIKREI 540
 481 ELPGAIVVLKKGKSKTEKEVMDYVAQVSNKRLRGVRFVDEVPKGLTGKIDKAIKREI 540
 541 LKKVAKM 548

Db 541 LKRPVAKM 548

RESULT 12

US-07-903-047-2

Sequence 2, Application US/07903047

Patent No. 529285

GENERAL INFORMATION:

APPLICANT: Kajiya, Naoki

APPLICANT: Nakano, Eiichi

TITLE OF INVENTION: Thermostable Luciferase Gene of Firefly,

TITLE OF INVENTION: DNA, And Process for The Preparation Of Thermostable

TITLE OF INVENTION: Luciferase Of Firefly

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036-2711

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/903,047

FILING DATE: 19920623

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Mistrock, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 7005-048

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 790-9090

TELEFAX: 212 869-8864/9741

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 548 amino acids

TYPE: AMINO ACID

STRANDEDNESS: single

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-07-903-047-2

Query Match 95.1%; Score 2684; DB 1; Length 548;

Best Local Similarity 93.4%; Pred. No. 5.5e-272;

Matches 512; Conservative 26; Mismatches 10; Indels 0; Gaps 0;

1 MENMENDENIVGPPPIEGSGAGAOILRYKMDRYAKLGAIAFTNALGVDTYAEYLE 60
 1 MENMENDENIVGPPPIEGSGAGAOILRYKMDRYAKLGAIAFTNALGVDTYAEYLE 60
 1 MENMENDENIVGPPPIEGSGAGAOILRYKMDRYAKLGAIAFTNALGVDTYAEYLE 60
 61 KSCCLGKALQNYGLVVDGRIALCSNCEFFPIVLAGLFIGVAVPNEIYTLRLVHSL 120
 61 KSCCLGKALQNYGLVVDGRIALCSNCEFFPIVLAGLFIGVAVPNEIYTLRLVHSL 120
 61 KSCCLGKALQNYGLVVDGRIALCSNCEFFPIVLAGLFIGVAVPNEIYTLRLVHSL 120
 121 GSKPTIVSSKGLDKYTVQKTYATATKYIYILSKYDYGOSMDNPIKNTPOGFKG 180
 121 GSKPTIVSSKGLDKYTVQKTYATATKYIYILSKYDYGOSMDNPIKNTPOGFKG 180
 121 GSKPTIVSSKGLDKYTVQKTYATATKYIYILSKYDYGOSMDNPIKNTPOGFKG 180
 181 SSFKTYEVNRKQVALINMSSGSLPGVQLTHENTVTRFSHARDPIYGNQVSPGTAVL 240
 181 SSFKTYEVNRKQVALINMSSGSLPGVQLTHENTVTRFSHARDPIYGNQVSPGTAVL 240
 181 SSFKTYEVNRKQVALINMSSGSLPGVQLTHENTVTRFSHARDPIYGNQVSPGTAVL 240
 241 TVVPRHNGGMPITGLGILCGFRVYMLKREDETFKLTDYKCTSVILVPTLAILNKS 300
 241 TVVPRHNGGMPITGLGILCGFRVYMLKREDETFKLTDYKCTSVILVPTLAILNKS 300

301 ELLDKYDLSNIVEIASGAPLSKEIGEAVARBNLPGRVGGTLETTSAITIIPEGDDK 360
 301 ELLDKYDLSNIVEIASGAPLSKEIGEAVARBNLPGRVGGTLETTSAITIIPEGDDK 360
 301 ELLDKYDLSNIVEIASGAPLSKEIGEAVARBNLPGRVGGTLETTSAITIIPEGDDK 360
 361 PGASGVVPLFKAKVIDLDTFKKTLGPNRGEVYCVAPMLKGYDNPATREIIDEGL 420
 361 PGASGVVPLFKAKVIDLDTFKKTLGPNRGEVYCVAPMLKGYDNPATREIIDEGL 420
 361 PGASGVVPLFKAKVIDLDTFKKTLGPNRGEVYCVAPMLKGYDNPATREIIDEGL 420
 421 HNGDIDGYDEEKNHFLVDRKLKIKYQVPPAELESYLQHPNIFDAGVAGVDPDPA 480
 421 HNGDIDGYDEEKNHFLVDRKLKIKYQVPPAELESYLQHPNIFDAGVAGVDPDPA 480
 421 HNGDIDGYDEEKNHFLVDRKLKIKYQVPPAELESYLQHPNIFDAGVAGVDPDPA 480
 481 ELPGAVVPLKKGKSMTEKVMYVAVSNAKRLRGVAFVDEVPKGLTGKIDKAIREI 540
 481 ELPGAVVPLKKGKSMTEKVMYVAVSNAKRLRGVAFVDEVPKGLTGKIDKAIREI 540
 481 ELPGAVVPLKKGKSMTEKVMYVAVSNAKRLRGVAFVDEVPKGLTGKIDKAIREI 540
 541 LKRPVAKM 548
 541 LKRPVAKM 548

RESULT 13

US-08-076-042-2

Sequence 2, Application US/08076042

Patent No. 5330906

GENERAL INFORMATION:

APPLICANT: KAJIYAMA, NAOKI

APPLICANT: NAKANO, EIICHI

TITLE OF INVENTION: MUTANT LUCIFERASE OF A FIREFLY, MUTANT

TITLE OF INVENTION: LUCIFERASE GENES, NOVEL RECOMBINANT DNAs CONTAINING THE

TITLE OF INVENTION: GENES AND A METHOD OF PRODUCING MUTANT LUCIFERASE

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: PENNIE & EDMONDS

STREET: 1155 AVENUE OF THE AMERICAS

CITY: NEW YORK

STATE: N.Y.

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/076,042

FILING DATE: 15-JUN-1993

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/675,211

FILING DATE: 26-MAR-1991

NAME: 435

ATTORNEY/AGENT INFORMATION:

NAME: MISTROCK, S. LESLIE

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 7005-026-999

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-790-9090

TELEFAX: 212-869-9741

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 548 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

ORGANISM: Luciola cruciata

US-08-076-042-2

Query Match 95.1%; Score 2684; DB 1; Length 548;

Best Local Similarity 93.4%; Pred. No. 5.5e-272;
Matches 512; Conservative 26; Mismatches 10; Indels 0; Gaps 0;

```

QY 1 MEMMENDENIYVGPPEPPIIEEGSAGAOIARKYMDRYAKIGAIAFTNALGVDVDTYAEYLE 60
Db 1 MEMMENDENIYVGPPEPPIIEEGSAGAOIARKYMDRYAKIGAIAFTNALGVDVDTYAEYLE 60
QY 61 KSCCLGKALONGGLVVDGRIALCSENCEEFFIPLVAGLFIVGVAFTNEITYLRELHVS 120
Db 61 KSCCLGKALONGGLVVDGRIALCSENCEEFFIPLVAGLFIVGVAFTNEITYLRELHVS 120
QY 121 GISKPTIVSSKKGDLKVTYVQKTAIKTIYILDSKVDYRGYQSDNFIKNTPOGFGK 180
Db 121 GISKPTIVSSKKGDLKVTYVQKTAIKTIYILDSKVDYRGYQSDNFIKNTPOGFGK 180
QY 181 SSFKTEVNRKKEQVALINMSSSGTGLPKGVOLTHENITVRFSHARDPIYGNVSGTAVL 240
Db 181 SSFKTEVNRKKEQVALINMSSSGTGLPKGVOLTHENITVRFSHARDPIYGNVSGTAVL 240
QY 241 TVVPHHGFMTTLLGYLTCGFRIVMLTKFDEETFLKTLQDYKCSVILVPTLFAILNRS 300
Db 241 TVVPHHGFMTTLLGYLTCGFRIVMLTKFDEETFLKTLQDYKCSVILVPTLFAILNRS 300
QY 301 ELLDKYDLSNIVEIASGAPLSKEIGEAARFNLPGVROGYGLTETTSAILITPEGDDK 360
Db 301 ELLDKYDLSNIVEIASGAPLSKEIGEAARFNLPGVROGYGLTETTSAILITPEGDDK 360
QY 361 PGASGRVVPLEFRAKYIDLDTKKTLGPNRRGEVGVKPMIMKGYVNDPEATREIIDEQWL 420
Db 361 PGASGRVVPLEFRAKYIDLDTKKTLGPNRRGEVGVKPMIMKGYVNDPEATREIIDEQWL 420
QY 421 HTGDIGYDEEHFIVDRKSLIKYKQVPPALESVLLQHPNIFDAGVAGVDPPIAG 480
Db 421 HTGDIGYDEEHFIVDRKSLIKYKQVPPALESVLLQHPNIFDAGVAGVDPPIAG 480
QY 481 ELPGAIVVLLKKGSMTEKEVMDYASQVSNARKRLRGVRFVDEVKGLTGKIDGRAIREI 540
Db 481 ELPGAIVVLLKKGSMTEKEVMDYASQVSNARKRLRGVRFVDEVKGLTGKIDGRAIREI 540
QY 541 LKKPVAKM 548
Db 541 LKKPVAKM 548

```

RESULT 14

US-09-380-061B-14
Sequence 14, Application US/09380061B
Patent No. 6265177

GENERAL INFORMATION:

APPLICANT: SQUIRRELL, DAVID JAMES
WHITE, PETER JOHN
LOWE, CHRISTOPHER ROBIN
MURRAY, JAMES AUGUSTUS HENRY
TITLE OF INVENTION: ENZYME ASSAY FOR MUTANT FIREFLY LUCIFERASE
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P. C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/380,061B
FILING DATE: 25-Aug-1999
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB98/01026

FILING DATE: 7-APR-1998
APPLICATION NUMBER: GB 9707468.8
FILING DATE: 11-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B. J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 124-725
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4000
TELEFAX: (703)816-4100
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-380-061B-14

Query Match

Best Local Similarity 95.1%; Score 2684; DB 4; Length 548;
Matches 512; Conservative 26; Mismatches 10; Indels 0; Gaps 0;

```

QY 1 MEMMENDENIYVGPPEPPIIEEGSAGAOIARKYMDRYAKIGAIAFTNALGVDVDTYAEYLE 60
Db 1 MEMMENDENIYVGPPEPPIIEEGSAGAOIARKYMDRYAKIGAIAFTNALGVDVDTYAEYLE 60
QY 61 KSCCLGKALONGGLVVDGRIALCSENCEEFFIPLVAGLFIVGVAFTNEITYLRELHVS 120
Db 61 KSCCLGKALONGGLVVDGRIALCSENCEEFFIPLVAGLFIVGVAFTNEITYLRELHVS 120
QY 121 GISKPTIVSSKKGDLKVTYVQKTAIKTIYILDSKVDYRGYQSDNFIKNTPOGFGK 180
Db 121 GISKPTIVSSKKGDLKVTYVQKTAIKTIYILDSKVDYRGYQSDNFIKNTPOGFGK 180
QY 181 SSFKTEVNRKKEQVALINMSSSGTGLPKGVOLTHENITVRFSHARDPIYGNVSGTAVL 240
Db 181 SSFKTEVNRKKEQVALINMSSSGTGLPKGVOLTHENITVRFSHARDPIYGNVSGTAVL 240
QY 241 TVVPHHGFMTTLLGYLTCGFRIVMLTKFDEETFLKTLQDYKCSVILVPTLFAILNRS 300
Db 241 TVVPHHGFMTTLLGYLTCGFRIVMLTKFDEETFLKTLQDYKCSVILVPTLFAILNRS 300
QY 301 ELLDKYDLSNIVEIASGAPLSKEIGEAARFNLPGVROGYGLTETTSAILITPEGDDK 360
Db 301 ELLDKYDLSNIVEIASGAPLSKEIGEAARFNLPGVROGYGLTETTSAILITPEGDDK 360
QY 361 PGASGRVVPLEFRAKYIDLDTKKTLGPNRRGEVGVKPMIMKGYVNDPEATREIIDEQWL 420
Db 361 PGASGRVVPLEFRAKYIDLDTKKTLGPNRRGEVGVKPMIMKGYVNDPEATREIIDEQWL 420
QY 421 HTGDIGYDEEHFIVDRKSLIKYKQVPPALESVLLQHPNIFDAGVAGVDPPIAG 480
Db 421 HTGDIGYDEEHFIVDRKSLIKYKQVPPALESVLLQHPNIFDAGVAGVDPPIAG 480
QY 481 ELPGAIVVLLKKGSMTEKEVMDYASQVSNARKRLRGVRFVDEVKGLTGKIDGRAIREI 540
Db 481 ELPGAIVVLLKKGSMTEKEVMDYASQVSNARKRLRGVRFVDEVKGLTGKIDGRAIREI 540
QY 541 LKKPVAKM 548
Db 541 LKKPVAKM 548

```

RESULT 15

US-09-111-752-7
Sequence 7, Application US/09111752
Patent No. 6074859

GENERAL INFORMATION:

APPLICANT: HIROKAWA, KOZO
APPLICANT: KAJIYAMA, NAOKI
APPLICANT: MURAKAMI, SEIJI

TITLE OF INVENTION: MUTANT-TYPE BIOLUMINESCENT PROTEIN, AND
TITLE OF INVENTION: PROCESS FOR PRODUCING MUTANT-TYPE BIOLUMINESCENT PROTEIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.,
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/111,752
FILING DATE: 08-JUL-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ. ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 552 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Luciola cruciata and Phantus pyralis
US-09-111-752-7

Query Match 91.7%; Score 2590; DB 3; Length 552;

Best Local Similarity 91.0%; Pred. No. 3.9e-262;
Matches 494; Conservative 31; Mismatches 18; Indels 0; Gaps 0;

QY 1 MNNNDENIVYGPPEPPYPIEGSSAGAOQLRRKYMDRYAKLGAIAFTNALTGVDYTYAEYLE 60
DB 1 MNNNDENIVYGPPEPPYPIEGSSAGAOQLRRKYMDRYAKLGAIAFTNALTGVDYTYAEYLE 60
QY 61 KSCCLGKALONGLYVDGRIALCSECEFFIPIAGLFIGVGAAPNEIYTLRELIVHSL 120
DB 61 KSCCLGKALONGLYVDGRIALCSECEFFIPIAGLFIGVGAAPNEIYTLRELIVHSL 120
QY 121 GISKPTIVFSSKGGIDVITVOKTVTAIKTIVILDSVDYRGYQSMNFKKNTPOGPKG 180
DB 121 GISKPTIVFSSKGGIDVITVOKTVTAIKTIVILDSVDYRGYQSMNFKKNTPOGPKG 180
QY 181 SSFRTVEVNRKEQVALIMSSGSTGLPKGVQLTHENIVTRFSHARDPIYGNQVSPGTAIL 240
DB 181 SSFRTVEVNRKEQVALIMSSGSTGLPKGVQLTHENIVTRFSHARDPIYGNQVSPGTAIL 240
QY 241 TVVPFHGFGMTTGLTGCGFRIVMLTKPDETEFLKTLQDYKSSVILVPTLFAIINRS 300
DB 241 TVVPFHGFGMTTGLTGCGFRIVMLTKPDETEFLKTLQDYKSSVILVPTLFAIINRS 300
QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGYGLTETTSAILITPEGDDK 360
DB 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGYGLTETTSAILITPEGDDK 360
QY 361 PGASGVVPLFFAKAVIDLDTKKTLLGPNRGEVCVKPMLMKGYVNDPEATREIIDEQWL 420
DB 361 PGASGVVPLFFAKAVIDLDTKKTLLGPNRGEVCVKPMLMKGYVNDPEATREIIDEQWL 420
QY 421 HTGDIGYDEEKHFFIYDRLKSLIKYGYVPAELESVLLQHPNIPDAGVAGLPDDPDIAG 480
DB 421 HTGDIGYDEEKHFFIYDRLKSLIKYGYVPAELESVLLQHPNIPDAGVAGLPDDPDIAG 480

Db 421 HTGDIGYDEEKHFFIYDRLKSLIKYGYVPAELESVLLQHPNIPDAGVAGLPDDPDIAG 480
QY 481 ELLPAVNVVLLKKGKSMTEKEVMDYVASQVSNKRLGSGVRFVDEYVKGTLGKIDGKAIRI 540
Db 481 ELLPAVNVVLEHGTWTEKEIVDYVASQVTTAKKLGGVVFDEYVKGTLGKIDGKAIRI 540
QY 541 LKK 543
Db 541 LKK 543

Search completed: November 29, 2001, 01:53:48
Job time: 3208 sec


```

|||||
101 ATGATATGCAAACTGGAGCAATTCCTTTACTAAGCACTTACCGGT 150
51 ValAspThrThrThrAlaGluTyrLeuGluLysSerCysLeuGly 67
|||||
151 GTGATATATAGTACGCGAATCTTAGAAAAATCATGCTGTCTAGAGA 200
67 ValAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
|||||
201 GGCTTTAAAGAAATTATGGTTGGTTGGTTGGATGGAAGAAATTCGGTATGCA 250
84 eacLusEncySGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
|||||
251 GTGAAACCTGTGAAGAAATCTTATTCCTGATATAGCCGCTTATTTATA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
|||||
301 GGCTGGGTGGTGGCTCCAACTAATGAGATTTACACTCTACGTCGATTTGGT 350
117 HisSerLeuGlyIleSerLysProThrIleValPheSerSerLysLysG 134
|||||
351 TCACAGTTTAGCATCTCTTAAGCCAAACAATTTATTAGTTCTTAAAAAG 400
134 LysLeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
|||||
401 GATTACATTAAGTTATTAAGCTACAAAAACGGTAACGCTATTTAAACC 450
151 IleValIleLeuAspSerLysValAspTyrArgIlyTyrGlnSerMetAs 167
|||||
451 ATTGTTATATTTGGACAGCAAGTGATTTAGAGTTATCAATCAATCGA 500
167 PheAsnThrIleLysLysAsnThrProGlnGlyPheLysGlySerPheI 184
|||||
501 CAACCTTATTTAAAAACACCTCCACAGATTTCAAAGGATCAAGTTTAA 550
184 ystThrValGluValAsnArgLysGluGluValAlaLeuIleMetLysSer 200
|||||
551 AAACGTGTAAGTTATACCCGAAAGAACAGCTTCTTTATATGAACCTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnI 217
|||||
601 TCGGGTTCACCGGTTTGGCCAAAGGTGTGCAACTACATCAATGAATTT 650
217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
|||||
651 GGTGATCAGATTTTCTCAGCTAGAGATCCAAATTTATGGAACCAAGTTT 700
234 ePProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
|||||
701 CACCAAGGACGGCTATTTTAACTAGTACCATTCATCATGTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetIle 267
|||||
751 ATGTTTACTACTTTAGGCTATCTACTGTTGTTTTCGATTTGTCATGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
|||||
801 AACGAATTTGACGAAGAGACTTTTAAAAACACTGCAAGATTTACAAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
|||||
851 GTTCAAGGTTATCTGTACCGACTTGTGTTGCAATTTCTTAATGAAAT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
|||||
901 GAATTACTGATAAATATGATTTATCAATTTAGTTGAAATTCATCTGG 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheA 334
|||||
951 CGGAGCACCCTTTATCTAAAGAAATGGTGAACCTTCTGTAACGTTTAA 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrSerAla 350
|||||

```

```

1001 ATTACGGGGTTCGTCAAGGCTATGCTTTACAGAAACACCTCTGCA 1050
351 IleIleIleThrProGluGluAspAspLysProGlyValAsnGlyLysVal 367
|||||
1051 ATTATTTATCACACCGAAGCGATGATAAACACAGCTCTTCTGGCAAGT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
|||||
1101 TGTGCCATTTATTTAAAGCAAAAGTTATCGATCTTGTATCTAAAAAACTT 1150
384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
|||||
1151 TGGGCCCGGACAGACGTCGAGAAAGTTGTGTAAAGGCTCTATAGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGlu 417
|||||
1201 AAAGGTTATGTAGATATATCCAGAACCAAGCAAGAAATCATAGATGAA 1250
417 uGlyTyrLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
|||||
1251 AGGTGGTTTCACACAGAGATATTGGGTATTACGATGAAGAAACAACTT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
|||||
1301 TCTTTATCGGTGATCGTTTGAAGCTTTTATCAAAATACAAAGATATCAA 1350
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
|||||
1351 GTACCACTCTGCTAATTTAGATCTGTTCTTTGCAACATCCAAATATTTT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyLysLeuProG 484
|||||
1401 TGATGCCGGGCTTGGCGGCTTCACATCTATAGCTGGAGACTTCGCG 1450
484 lValAlaValValIleLysLysGlyLysSerMetThrGluLysGluVal 500
|||||
1451 GAGCTGTGTTGTACTTAAAGAAAGAAATCTATGACTGAAAAAGAACTTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgLys 517
|||||
1501 ATGGATTAGCTTCTAGTCAAGTTTCAAAATGCCAAACGTTTGGGTGGTGG 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
|||||
1551 TGTCCGTTTGTGGAGCAAGTACTAAAGTCTCACTGTTAAATGAGCG 1600
534 lLysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
|||||
1601 GTAAAGCAATTTAGAGAAATCTGAAAGAAACAGTTGCTTAAGATG 1644

```

seq_name: /cgn2_6/plodata/2/lna/5A_COMB.seq:05-07-903-047-7

seq_documentation_block:

Sequence 7, Application US/07903047

Patent No. 529285

GENERAL INFORMATION:

APPLICANT: KajiYama, Naoki

APPLICANT: Nakano, Eiichi

TITLE OF INVENTION: Thermostable luciferase of Firefly.

TITLE OF INVENTION: Thermostable luciferase Gene of Firefly, No. 529285el Reco

TITLE OF INVENTION: DNA, And Process For The Preparation Of Thermostable

TITLE OF INVENTION: Luciferase Of Firefly

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/903,047
 FILING DATE: 19920623
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Mistrock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 7005-048
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 790-9090
 TELEFAX: 212 869-8864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1644 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: cDNA to mRNA
 US-07-903-047-7

alignment_scores:
 Quality: 2814.00 Length: 548
 Ratio: 5.144 Gaps: 0
 Percent Similarity: 99.818 Percent Identity: 99.635

alignment block:
 US-09-581-241-6 x US-07-903-047-7 ..

Align seg 1/1 to: US-07-903-047-7 from: 1 to: 1644

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProph 17
 1 ATGGAACAAATGAGACGATGAAATATGCTGATGCTGCTGAACCATTT 50
 17 eTyrProIleGluGluGlySerAlaGlyAlaGluLeuArgLysTyrMet 34
 51 TTACCTATTTGAAGAGGATGCTGAGACACAAATTCGCAAGTATATGG 100
 34 sPArGlyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
 101 ATCGATATCAAACTTGAGCAATTCCTTTACTAACGCACATTCACCGGT 150
 51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysLeuGlyGcl 67
 151 GTCGATTATACGTACGCCAATACCTAGAAAATCATGCTGTCTAGAGA 200
 67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyrAlaIleAlaLeuCys 84
 201 GGTCTTAAAGAAATATGTTGTTGTTGATGAGAAATTCGCTATGCA 250
 84 eGluAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
 251 GTGAAAACGTGAGAAATCTTATCTCTGTAATAGCCGCTTATTTATA 300
 101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeu 117
 301 GGGTGCGGTGCTCCAACTAATGATTTACACTCTACGTGAATTTGGT 350
 117 IHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysGly 134
 351 TCACAGTTTAGCATCTCTAAGCAACATTTGATTACTTCTAAAAAAG 400
 134 ILeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
 401 GATTAGTAAGTTAATACGTACAAAAACGTAACCTATTAATAAAC 450
 151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
 451 ATTGTATATTTGACAGCAAGTGTATATAGGTTATCAATCCATCGGA 500

167 pAsnPheIleLysAsnThrProGlnGlyPheLysGlySerSerPheIle 184
 501 CAACCTTATTAATAAAAAACACCTCCACAGGTTTCAAGGATCAAGTTTA 550
 184 yStrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
 551 AAACGTAGAGTTAACCCGAAAGAACAGTTGCTCTTATTAATGAACCTCT 600
 201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnIle 217
 601 TCGGGTTCACACCGTTTGCCAAAAGCTGTGCAACTTACATCAAAAAATGC 650
 217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnVal 234
 234 ePProGlyThrAlaIleLeuThrValValProPheHisGlyPheGly 250
 701 CACAGGACAGCGCTATTTTAAGTGTAGTACCATTCATCATGCTTTGGT 750
 251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgGlyValMetIle 267
 751 ATGTTTACTACTTGAAGCTATCTAAGTGTGTTTGTGATTTGTCATGTT 800
 267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
 801 AACCAATTTGACGACAGACACTTTTAAAAACACGCAAGATACAAAT 850
 851 GTTCACAGCTTATCTTGTACCGACCTTGTTCATTTCTTAATTAAGAT 900
 301 GluLeuAspLysTyrAspLeuSerAsnLeuValGluIleLasergl 317
 901 GAATTAATCGATTAATATGATTTATCAAAATTTAGTCAAAATTCATCTGG 950
 317 yGlyAlaProLeuSerLysGluIleGlyAlaValAlaArgArgPhe 334
 951 CGGAGACCTTATCTAAGAAATGCTGAAGCTGTGTACAGCTTTTA 1000
 334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
 1001 ATTTACCGGGTGTCTCAAGCTATGCTTTAACAGAAACACCTCTGCA 1050
 351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVal 367
 1051 ATTATTTATCACCCGAAAGCGATGATTAACACAGCTGCTCTGCAAAAGT 1100
 367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
 1101 TGTGCAATTTATTAAGCAAAAGTTATCGATCTTGATCAAAAAAACTT 1150
 384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
 1151 TGGGCCGACAGACGTGAGAGATTTGTGAAGGCTTCATATCTTATG 1200
 401 LysGlyTyrValAspAsnProGluAlaThrArgLysIleIleAspGlu 417
 1201 AAAGTTATGTAGATTAATCCAGAAAGCAACAGAAATCATAGATGAAGA 1250
 417 uGlyThrLeuHisThrGlyAspIleGlyTyrTyrAspGluLysHisP 434
 1251 AGGTGTTGTCACACAGAGATATTGGTATTTACAGTAAGAAAAACATTT 1300
 434 hepPheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
 1301 TCTTATCTGTCGATGCTTTGAAGTCTTAAATCAATAAAGAGATACCAA 1350
 451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePhe 467
 1351 GTACACCTGCTGAATAGAAATCTGTTCTTTGCAACATCCAAATATATTT 1400

467 easpaalaglyvalaaglyvalproasprolielaglygylureu prog 484
 1401 TGAATGCCGGCTGCTGCGTCCAGATCCTATATGCTGGAGCTTCGG 1450
 484 lyalaValaValleuLysGlyLysSerMetThrGlyLysGlyVal 500
 1451 GAGCTGTGTTGACTTGAAGAAAGAAATCTATGACTGAAAAAGAGTA 1500
 501 MetAspTyrAlaAlaSerGlnValSerAlaLysArgLeuArgGly 517
 1501 ATGATTACTGCTGCTAGTCAAGTTCAATGCAAAACGTTGCGTGGTGG 1550
 517 yValaArgpHeValaSpGluValProLysGlyLeuThrGlyLysIleAsp 534
 1551 TGTCCGTTTGTGACGAAAGTAAAGCTCACTGCTAAATGACG 1600
 534 lyysalalaleargjuilleuLysLysProvalAlaLysMet 548
 1601 GTAAACCAATTAGAGAAATCTGAAAGAACCAAGTGTCTAAGATG 1644

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:us-09-380-061B-15

seq_documentation_block:

Sequence 15, Application US/09380061B
 Patent No. 6265177

GENERAL INFORMATION:

APPLICANT: SOUIRRELL, DAVID JAMES

WHITE, PETER JOHN

LOWE, CHRISTOPHER ROBIN

MURRAY, JAMES AUGUSTUS HENRY

TITLE OF INVENTION: ENZYME ASSAY FOR MUTANT FIREFLY LUCIFERASE

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESS: NIXON & VANDERHUYE P.C.

STREET: 1100 NORTH GLEBE ROAD

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: U.S.A.

ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/380.061B

FILING DATE: 25-Aug-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB98/01026

FILING DATE: 7-APR-1998

APPLICATION NUMBER: GB 9707468.8

FILING DATE: 11-APR-1997

ATTORNEY/AGENT INFORMATION:

NAME: SADOFF, B. J.

REGISTRATION NUMBER: 36,663

REFERENCE/DOCKET NUMBER: 124-725

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)816-4100

TELEFAX: (703)816-4100

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 1644 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

MOLECULE TYPE: cdna

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1644

SEQUENCE DESCRIPTION: SEQ ID NO: 15:

US-09-380-061B-15

alignment_scores:

Quality: 2814.00

Ratio: 5.144

Percent Similarity: 99.818

Percent Identity: 99.635

alignment_block:

US-09-581-241-6 x US-09-380-061B-15

Align seq 1/1 to: us-09-380-061B-15 from: 1 to: 1644

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProph 17
 1 ATGGAAGAAACATGGAAGACGATGAAATATTTGATGATGCTCTGAACCAT 50
 17 eTyrProIleGluGluGlySerAlaGlyAlaGlnLeuArgLysTyrMet 34
 51 TTACCCCTATTGAAGAGGATCTGCTGGAGCACAATTGCCAGATATATGG 100
 34 sPArgTyrAlaLysIleuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
 101 ATCGATATGCAAAACCTTGAGCAATTCCTTACTTAACGCACTTACCGGT 150
 51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGly 67
 151 GTGATTTATACGTACCGCGAATCTTGAAGAAATCATGCTGCTTAGAGCA 200
 67 uAlaLeuLysAsnTyrGlyLeuValAlaSpGlyArgGlyAlaLeuCys 84
 201 GGCTTTAAAGAAATTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 250
 84 erGluAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
 251 GTGAAACCTGTGAAGAAATCTTATTCCTGTTATTAACCGGTTATTTATA 300
 101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
 301 GGTCGTGCTGTGCTCCAACTATGAGATTTCACCTACGTAATTTGGT 350
 117 HisSerLeuGlyIleSerLysProThrIleValPheSerSerLysGly 134
 351 TCACACTTATGAGCATCTTACGCAACAATGTAATTTACTTTAAAGAAAG 400
 134 lyLeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
 401 GATTAGATAAGTTATTAAGTACAAAAAAGCGTACTCTTTTAAAGC 450
 151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
 451 ATTGTTATATTGACAGCAAGTGATTTAGAGTTATCAATCCATGCA 500
 167 PAsnPheIleLysLysAsnThrProGlnIlePheLysGlySerSerPhe 184
 501 CAACCTTATTAATAAAAAACACCTCCACAAAGTTTCAAGGATCAAGTTTA 550
 184 ySerThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
 551 AAACGTGAGAAAGTTAAACGCAAGAAAGAGTTGCTTTATATGAACTGT 600
 201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsn 217
 601 TCGGTTCAACCGGTTTGCAGAAAGGTGTGCAACTTACTCAAGAAATGTC 650
 217 eValThrArgPheSerHisAlaArgProIleTyrGlyAsnGlnVal 234
 651 AGTACTAGATTTTCTACGCTAGAGATCCAAATTTATGAAACCAAGTTT 700
 234 erProGlyThrAlaIleLeuThrValProPheHisIleGlyPheGly 250
 701 CACGAGCAGCGGTATTTTAAGTATGATCATCATCATGATGTTTGGT 750
 251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheAlaGlyIleValMet 267

```

751 ATGTTTACTTACTTAGGCTATCTACTTGTGTTTGTGTAATTCATGCTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrIle 284
801 AACGAAATTTGACGAGACACTTTTAAAAACACTGCAGATACCAAT 850
284 ySsrSerValIleLeuValProThrLeuPheAlaIleuAsnArgSer 300
851 GTTCAAGCGTATCTTCTTACCAGCTTGTGTCATTCCTTAATAGAGT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSer 317
901 GAATTACCGATTAATATGATTTTCAAAATTTAGTGAATTCATCGTCG 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArghea 334
951 CGGAGCACCCTTATCTAAAGAAATGGGAGCGTTCGCTAGACGCTTTA 1000
334 sPleuProGlyValArgGlnGlyThrGlyLeuThrGluThrSerAla 350
1001 ATTACCGGGTGTGTCAGGCTATGCTTAAACAGAAACACTCTGCA 1050
351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLys 367
1051 ATTATTATCACACCGAAGCGGATGATTAACAGAGTCTGCTGCAAAAGT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysThrL 384
1101 TGTGCAATTAATTTAAAGCAAAAGTTATCGATCTGATCAAAAAAATT 1150
384 euGlyProAsnArgArgLysGlyGluValLysValLysGlyProMetLeu 400
1151 TGGGCCCCACAGACGTGAGAGATTTGTAAAGGCGCTATGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleLeuAspGlu 417
1201 AAAGGTATGTAGTAATCCAGAACCAACAGAAATCATAGATGAAGA 1250
417 uGlyTrpLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysH 434
1251 AGCTGTGGTGCACACAGAGATTTGGGATTCATCATAGACAAAAACATT 1300
434 hepHeiIleValAspArgLeuLysSerLeuIleLysTyrLysGlyGln 450
1301 TCTTATCTGATGATGCTTGAAGCTTTAATCAATACAAAGATATCAA 1350
451 ValProProAlaGluLeuGlnSerValLeuLeuGlnHisProAsnIle 467
1351 GTACCACTGCTGAATAGAAATCTGCTTTGCAACATCCAAATATTTT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuPro 484
1401 TGTATGCCGGCGTGTGCTGGCGTTCAGATCTTATAGCTGTGAGCTTCC 1450
484 lValAlaValAlaValLeuLysLysGlyLysSerMetThrGlnLysGlu 500
1451 GAGCTGTGTTGTAAGTGAAGAAATCTATGACTGCAAAAGAGTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
1501 ATGGATTAGCTGTGCTAGTCAGATTCCTCAAAATGCAAAACGTTTCG 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAsp 534
1551 TGTCCGTTTGTGACGAGAGTACTAAAGTCTCACTGTGTAATATGAGC 1600
534 lYlYsAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
1601 GTAAGCAATTAGAGAAATCTGAGAGAAACAGATTGCTAAGATG 1644

```

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq.us-08-460-934-8
seq_documentation_block:

```

: Sequence 8, Application US/08460934
: Patent No. 5814465
: GENERAL INFORMATION:
: APPLICANT: TATSUMI, HIROKI
: APPLICANT: FUKUDA, SATOSHI
: APPLICANT: KIKUCHI, MAMORU
: APPLICANT: KOBAYASHI, YASUJI
: TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
: TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
: TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT
: ANALYSTS METHOD
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: P. C.
: STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
: CITY: ARLINGTON
: STATE: VA
: COUNTRY: USA
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/460,934
: FILING DATE: 05-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 193798/1994
: FILING DATE: 27-JUL-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 54625/1995
: FILING DATE: 14-MAR-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 98857/1995
: FILING DATE: 24-APR-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: OBLON, NORMAN F.
: REGISTRATION NUMBER: 24,618
: REFERENCE/DOCKET NUMBER: 7126-001-0
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-413-3000
: TELEFAX: 703-413-2220
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1908 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: unknown
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: misc. feature
: LOCATION: 1..1908
: OTHER INFORMATION: /note="The nucleotide sequence of
: OTHER INFORMATION: the biotinylated firefly luciferase gene contained in
: OTHER INFORMATION: recombinant plasmid pHLf248 DNA"
: NAME/KEY: CDS
: LOCATION: 1..1908
: US-08-460-934-8

```

alignment_scores: Length: 547
Quality: 2812.00 Gaps: 0
Ratio: 5.141
Percent Similarity: 100.000 Percent Identity: 99.634

alignment_block:
US-09-581-241-6 x US-08-460-934-8 ..
Align seg 1/1 to: US-08-460-934-8 from: 1 to: 1908

```

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProph 17
1 ATCGAAACACGTGAGAAACGATGAAATATTGTGTATGCTTACCAACCAT 50
17 eYrProIleGlyGlySerAlaGlyAlaGluLeuArgLysTyrMet 34
51 TTACCTATTGAAGAGGATCTGCTGAGACCAATTGCCCAAGTATATGG 100
34 sParGTrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 ATCGATATGCAAACTTGGAGCATTTGCTTTACTAAGCATTTACCGGT 150
51 ValAspTyrThrTyrAlaGlyTyrLeuGlyLysSerCysCysLeuGly 67
151 GTGCAATTATACGACCGCAATCTAGAAAATCATGCTGCTAGAGAG 200
67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
201 GGCTTTAAAGAAATTATGTTGGTTGTTGATGGAAGAAATTCCTTATG 250
84 eRGluAsnCysGlyGluPhePheIleProValLeuAlaGlyLeuPheIle 100
251 GTGAAACACGTGAGAGTCTTATTCTGTATTTACCGGTTTATTTATA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeu 117
301 GGTTGCGGTGGCTCCAACTATGAGATTTACACTCTACGTAATTTGGT 350
117 HisSerLeuGlyIleSerLysProThrIleValPheSerLysLysG 134
351 TCACAGTTATAGCATCTTACGCCAACAAATGATTTAGTTCTTAAAG 400
134 ILeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
401 GATTTAGATTAAGTATTAAGTATACAAAAACGTAACCTCTATTAAAC 450
151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
451 ATTGTTATATTTGACAGCAAAAGTATTAAGTATTAAGTATTAAGTAT 500
167 PAsnPheIleLysLysAsnThrProGlnLysPheLysGlySerSerPhe 184
501 CAACCTTTATTAATAAAAAACACTCCAAAGGTTTCAAAAGATCAAGTT 550
184 YSThrValGluValAsnArgLysGluValAlaLeuIleMetAsnSer 200
551 AAACGTGTAAGTAAACCGCAAGAAAGAGTTGCTTTATATGAACCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsn 217
601 TCGGTTTACCGGTTTGCACAAAGGTTGCAACTTACTCATGAAATTT 650
217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnVal 234
651 GGTCACGCGGTTTCTCACGCTAGAGATCCAAATTTATGAAACCAAGTT 700
234 eRPProGlyThrAlaIleLeuThrValValProPheHisHisGlyPhe 250
701 CACCAAGGACGCTATTTTAACTGTAGTACCATTCATCATGCTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleAlaMet 267
751 ATGTTTACTTATAGCTATCTAATCTGTGTTTCTGATATGTATGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLys 284
801 AACGAAATTTGACGAAGAGACTTTTAAACACTGCACAGATTCACAA 850
284 YsSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTTCAAGGCTATTTCTGTACGACTTTGTTGCAATTTAATAGAGT 900

```

```

301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
901 GAATTTACTGCATAAATGATTTATCAAAATTTAGTTGAAATTCATCTGG 950
317 YGlyAlaProLeuSerLysGluIleGlyValAlaValAlaArgArPhe 334
951 CGGAGACCTTTATCTAAGAAATTTGTTGAAGCTGTTGCTAGACGTTT 1000
334 sLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrHisAla 350
1001 ATTTACCGGGTGTTCGTCACAGGCTATGTTTAAACAGAAACACCTGCA 1050
351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerLysVal 367
1051 ATTTATTTATCACCCGGAAGGCGATGATAAACAGGTCCTTCTGGCAAG 1100
367 ValProLeuPheLysAlaLysValIleAspLeuAspThrLysThr 384
1101 TGTGCATTTATTTAAAGCAAAAGTATTCGATCTTGATTAATAAAACCT 1150
384 eUGlyProAsnArgArgGlyGlyValCysValLysGlyProMetLeuMet 400
1151 TGGGCCCGAAGACGCTGGAGAGTTTGTGTAAGGCTCTATGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGlu 417
1201 AAAGTTATGATGATATATCCAGACGACACAGAGAAATCATGATGATA 1250
417 uGlyThrLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHis 434
1251 AGTTGGTTGCACACAGAGATTTGGTATTAAGTATTAAGTATTAAGT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
1301 TCTTTATCTGATCTTGAAGCTCTTTATGCAAAATCAAGAGATTTCA 1350
451 ValProProAlaGluLeuGlnSerValLeuLeuGlnHisProAsnIle 467
1351 GTACACCTGCTGAATTAAGATCTGTTTGGCAACATCCAAATTTT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuPro 484
1401 TGATGCGGCGTGTGCTGGCGTTCAGATCTTATAGCTGCTGAGCTCCG 1450
484 IValAlaValAlaLeuLysLysGlyLysSerMetThrGluLysGluVal 500
1451 GAGCTGTTGTTGCTGTAAGAAAGAAATCTATGACTAATAAAAGAGTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
1501 ATGATTTACGTTGCTAGTCAAGTTTCAATGCAAAACGTTTCTGCTGG 1550
517 YValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAsp 534
1551 TGTCCGTTTGTGGAGCAAGTACCTAAGGTCTCACTGCTGTAATAATG 1600
534 IYsAlaIleArgGluIleLeuLysLysProValAlaLys 547
1601 GTAAAGCAATTTAGAAATTAAGTGAAGAAACAGTTGCTAAG 1641
seq_name: /cgn2_6/ptcddata/2/ina/5B_COMB.seq:us-08-782-118-8
seq_documentation_block:
: Sequence 8, Application us/08782118
: Patent No. 5843746
: GENERAL INFORMATION:
: APPLICANT: TATSUMI, HIROKI
: APPLICANT: FUKUDA, SATOSHI
: APPLICANT: KIKUCHI, MAMORU
: APPLICANT: KOYAMA, YASUJI
: TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
: TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A GENE
: TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINATED AND A BIOLUMINESCENT

```

```

? TITLE OF INVENTION: ANALYSIS METHOD
? NUMBER OF SEQUENCES: 14
? CORRESPONDENCE ADDRESS:
? ADDRESS: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
? ADDRESSEE: P.C.
? STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
? CITY: ARLINGTON
? STATE: VA
? COUNTRY: USA
? ZIP: 22202
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/782,118
? FILING DATE: 13-JAN-1997
? CLASSIFICATION: 435
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 08/460,934
? FILING DATE: 05-JUN-1995
? APPLICATION NUMBER: JP 193798/1994
? FILING DATE: 27-JUL-1994
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: JP 54625/1995
? FILING DATE: 14-MAR-1995
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: JP 98857/1995
? FILING DATE: 24-APR-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: OBLON, NORMAN F.
? REGISTRATION NUMBER: 24,618
? REFERENCE/DOCKET NUMBER: 7126-001-0
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 703-413-3000
? TELEFAX: 703-413-2220
? INFORMATION FOR SRO ID NO: 8:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1908 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: unknown
? MOLECULE TYPE: DNA (genomic)
? FEATURE:
? NAME/KEY: misc.feature
? LOCATION: 1..1908
? OTHER INFORMATION: /note= "the nucleotide sequence of
? OTHER INFORMATION: the biotinylated firefly luciferase gene contained in
? OTHER INFORMATION: recombinant plasmid pHLf248 DNA"
? NAME/KEY: CDS
? LOCATION: 1..1908
? US-08-782-118-8

```

```

alignment_scores:
    Quality: 2812.00      Length: 547
    Ratio: 5.141         Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 99.634

```

```

alignment_block:
US-09-581-241-6 x US-08-782-118-8 ..
Align seg 1/1 to: US-08-782-118-8 from: 1 to: 1908

```

```

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProph 17
1 ATGGAAGACATGAGACGATGAAATATGTGATGCTGCAACCAATT 50
17 eTyrProIleGluGluGlySerAlaGlyAlaGlnLeuArgLysTyrMetA 34
51 TTACCTATTGGAAGAGATCTGCTGAGACAAATTGCCCAAGTATATGG 100

```

```

34 spArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 ATCGATATCCAAACTGGAGCAATTTGCTTTACTACGCACTTACCGGT 150
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGlyG 67
151 GTGGATTATACGTACGCCCAATCTAGAAAATCATCGTCTGTAGAGAG 200
67 ValAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeu 84
201 GGCTTTAAAGAAATTATGCTTGTGTTGATGGAAGAAATGGCTATGCA 250
84 eGluAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
251 GTGAAACTGTGAGAGTCTTATTCCTGATTTACCGGTTATTATTANA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeu 117
301 GTGTGCGGTGTGCTCCAACTATGAGATTACACTGCTGCAATTTGCT 350
117 ThrSerLeuGlyIleSerLysProThrIleValPheSerLysLysG 134
351 TCACAGTTTAGCATCTCTAAGCAACAATGTATTAGTTCTTAAAAAG 400
134 LysLeuAspLysValIleThrValGlnLysThrValThrAlaIleLys 150
401 GATTAGTAAAGTTATTAAGTCTCAAAAAACGGTAAGTCTATTAAGCC 450
151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
451 ATTGTATATTGACAGCAAGGATTAAGAGTTATCAATCAATGAG 500
167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPhe 184
501 CAAGTTATTAAAAAACAACCTCCACAAGTTTCAAAAGATCAAGTTTA 550
184 YsThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
551 AAAGTATGAGAGTTAACCCCAAGAACAGTTCCTTATTAAGCAACTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsn 217
601 TCGGTTCAACCGGTTGCCAAAAGGTGTCACACTTACTATGAAAATT 650
217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnVal 234
651 GGTACAGCGTTTCTCAGCGTACAGATCCATTTATGAAACCAAGTTT 700
234 eProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
701 CACAGCGACGCTATTTAAGTCTAGTACCATTCATCATGCTTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMet 267
751 ATGTTACTACTTAAAGCTATCTAAGTGTGTTGTTGCTATGTCAGTGT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLys 284
801 AACGAAATTTGACGAGAGACTTTTAAAAACACTCAGAGATTAACAAT 850
284 YsSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTTCAAGGTTATCTGTACCGACTTGTGTCATTTCTTAATAGAACT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
901 GAATTACTCGAATAATATGATTTATCAAAATTTAGTGAATTTGCATCGG 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
951 CGAGACACCTTATCTAAGAAATTTGGAAGCTGTTGCTGAGAGTTTAA 1000

```

```

334 snleuproglyvalarglnglytyrlyleuthrcluthrthrserala 350
1001 ATTACCGGGTTCGTCAGAGGCTATGCTTACAGAACACCTCTGCA 1050
351 llellellehrprogluglyaspaplysproglyalaserglylsva 367
1051 ATTATATATCACCCGAGAGGCGATGATTAACCGAGTCTTGGCAAGT 1100
367 ValProlePheLysAlaLysValIleaspLeuaspThrlylsvThL 384
1101 TGTGCATTATTAAAGCAAACTTTCATCTGTGATCTTAAACAACTT 1150
384 euglyproasnargarglygluvalcysvallylsvPrometleumet 400
1151 TGGGCCCGAACACGCTGAGAGTGTGTAAAGGCTCTTATGCTTANG 1200
401 LysGlyTyrValaspasnprogluValThrargGluIleleaspGlu 417
1201 AAAGTTATGTGATATATCCAGAGCAACAGAAATCATGATGAGACA 1250
417 uGlyThrLeuHisThrGlyAspIleGlyTyrTyrAspGluGlySHSP 434
1251 AGTTGTTGTCACACAGAGATATTGGTATTACGATGAGAAAAACATT 1300
434 hePheIleValaspargleuLysSerLeuIleLysTyrLysGlyTyrGln 450
1301 TCTTATCTGTGATCGTTGAAGTCTTATATCAATAACAAAGATATCAA 1350
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
1351 GTACACCGCTGATTAATCACTGTTCTTTGCAACATCCAAATATTTT 1400
467 easPalaglyValAlaGlyValProaspProIleAlaGlyLeuLeuprog 484
1401 TGATCCCGCGTTCGTCGCTCCAGATCTATAGCTGCTGAGCTTCGG 1450
484 lYAlaValAlaValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
1451 GAGCTGTTGTTGTTACTGAAAAAGAAATCTATGACTGAAAAAGAACTA 1500
501 MetaspTyrValAlaSerGluValSerAsnAlaLysArgLeuArgGly 517
1501 ATGATTTACGCTGTAGTCAAGTTCAATGCAAAACGTTTGGCGTGG 1550
517 yValArgPheValaspGluValProLysGlyLeuThrGlyLysIleaspG 534
1551 TGTCCGTTTGTGGAGAGTACCTAAAGGTCTCAGTGTAAATGTGACG 1600
534 lYlYsAlaIleArgGluIleLeuLysLysProValAlaLys 547
1601 GTAAAGCAATTAGAGAAATACTGAAGAAACAGTTGCTTAAG 1641
seq_name: /cgn2_6/prodata/2/ina/5a_COMB.seq:US-08-460-934-5
seq_documentation_block:
; Sequence 5, Application US/08460934
; Patent No. 5814465
; GENERAL INFORMATION:
; APPLICANT: TATSUMI, HIROKI
; APPLICANT: FUKUDA, SATOSHI
; APPLICANT: KIKUCHI, MAMORU
; APPLICANT: KOTAMA, YASUJI
; TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
; TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
; TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT
; TITLE OF INVENTION: ANALYSIS METHOD
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBION, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; STATE: VA

```

```

COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,934
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 193798/1994
FILING DATE: 27-JUL-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 54625/1995
FILING DATE: 14-MAR-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 98857/1995
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBION, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: unknown
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1704
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..1704
OTHER INFORMATION: /note="Nucleotide sequence of the
OTHER INFORMATION: biotinylated firefly luciferase gene contained in recombin
US-08-460-934-5
alignment_scores:
Quality: 2799.00 Length: 546
Ratio: 5.126 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.267
alignment_block:
US-09-581-241-6 x US-08-460-934-5 ..
Align seg 1/1 to: US-08-460-934-5 from: 1 to: 1704
3 AsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPheTyrPr 19
67 AGCTCGAGACGATGAAATATTTGTGATGCTCTGAAACCACTTTTACC 116
19 cllleglugluglyseralaglyalaglnleuarglystyrmetaspargt 36
117 TATGTAAGAGGAGATCGTGGAGCAAAATTCGCAAGTATATGATCAT 166
36 yValAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGlyValasp 52
167 ATGCAGAACTTGAGCAATGCTTTTACTTAACGCACTTACCGGTGCGAT 216
53 TyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGlyGluAlaLe 69
217 TATAGTACCGCGCAATACTTAGAAAAATCATGCTGTAGAGAGAGCTTT 266
69 ulysasnTyrGlyLeuValValaspGlyArgIleAlaLeuLysSerGln 86

```

```

1167 ATTATTAAGCAAAAGTTATCGACTTGTATCTAATAAAGAACTTGGGCC 1216
386 roasnarqarglygluvalcysvalylslygylprometleumelysgly 402
1217 CGACAGACAGCTGGAGAGTTGTGTAAGGGCTCTATCTTATGAAGGT 1266
403 tyvalaspsanprogluathrarqgluilelleaspgluglytyr 419
1267 TATGTAGATTAATCCAGAACCAAGCAAGAAATCATAGATGAAGAAGTTG 1316
419 pleuhsthrqlyaspilleglytyrtyraspgluglyulysshisphet 436
1317 GTTGACACAGAGAGATATGGTATACGATGAAGAAAAACATTTCTTTA 1366
436 levalasprgleuysserleuilelystlysglytyrslvalpro 452
1367 TCGTGATCGTTGAAGCTTTTAATCAATACAAAGATATCAATACCA 1416
453 proalagluleugluservalleuuglnhlsproasnlelphespa 469
1417 CCGCTGAATTAAGATCTGTTCTTTGCAACATCCAAATATTTTGATGC 1466
469 aglyvalalaglyvalproasproilleaqlglyleuproglyalav 486
1467 CGCGCTTGCTGGCGTCCAGATCTATAGCTGTGAGCTTCGGAGCTG 1516
486 alvalvalleuyslysglytyrsermetthrqlulysgluvalmetasp 502
1517 TTGTTTACTGTAANAAGAAATCTATGCTGCAAAAAAGATTAATGAT 1566
503 tyvalalaserqinalserasnalaalysargleuarglyglyvalar 519
1567 TACGTTGTAGTCAAGTTCAATGCAAAAACGTTGCGTGCGTCCG 1616
519 gphvealaspgluvalprolysglyleuthrqllysleaspglylva 536
1617 TTTTGGAGAGAGTCTTAAGGTCTCACTGTAATTAATTAAGGATTAAG 1666
536 lalleargluileuyslyspovalalalysmet 548
1667 CAATTAGAGAAATCTGAAGAACCAAGTTGCTTAAGATG 1704

```

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-782-118-5

seq_documentation_block:

Sequence 5, Application us/08782118

Patent No. 5843746

GENERAL INFORMATION:

APPLICANT: TATSUMI, HIROKI

APPLICANT: FUKUDA, SATOSHI

APPLICANT: KIKUCHI, MAMORU

APPLICANT: KOYAMA, YASUJI

TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE

TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A

TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT

TITLE OF INVENTION: ANALYSIS METHOD

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESSES:

ADDRESSEE: P. C. MCCLELLAND, MATER & NEUSTADT,

STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR

CITY: ARLINGTON

STATE: VA

COUNTRY: USA

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/782,118

FILING DATE: 13-JAN-1997

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/460,934
FILING DATE: 05-JUN-1995
APPLICATION NUMBER: JP 193798/1994
FILING DATE: 27-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 54625/1995
FILING DATE: 14-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 98857/1995
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1704
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..1704
OTHER INFORMATION: /note= "Nucleotide sequence of the
OTHER INFORMATION: biotinylated firefly luciferase gene contained in recombinant
OTHER INFORMATION: plasmid pHLf203 DNA"
US-08-782-118-5

alignment_scores:

Quality: 2799.00	Length: 546
Ratio: 5.126	Gaps: 0
Percent Similarity: 100.000	Percent Identity: 99.267

alignment_block:

US-09-581-241-6 x US-08-782-118-5 ..

Align seg 1/1 to: US-08-782-118-5 from: 1 to: 1704

```
3 AsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPheTyrPr 19
:.....
67 AGTCGCGAAGCATGAAATATGTGTATGCTGCTGACCATTTTACCC 116
19 oiliGluGluGlySerAlaGlyAlaGluLeuArgLysTyrMetAspArgT 36
|||||
117 TATGTAAGAGGAGATCTGCTGAGCACAATTGCCAAGTATATGATGAT 166
36 YrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGlyValAsp 52
|||||
167 ATGCAAACTTGGAGCAATTGCTTTTACTAACGACACTTACCGGTGAT 216
53 TyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGlyAlaLe 69
|||||
217 TATACGTACCGCGAATCTTAGAAAATCATGCTGTAGAGAGCGCTT 266
69 uLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCysSerGlu 86
|||||
267 AAGGAATTATGTTGGTGTGTGATGGAAGAAATGCGTTATGCAAGTAAA 316
86 snCysGluGluPhePheIleProValLeuAlaGlyLeuPheIleGlyVal 102
|||||
317 ACAGTGAAGGTTCTTATTCCTGATATACCGGTTTATTTAAGGTGTC 366
103 GlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuValHis 119
```

```
|||||
367 GGTGGCTCCAACTAATGAGATTATACCTCTACGTAATGGTTCACAG 416
119 rLeuGlyIleSerLysProThrIleValPheSerSerLysGlyLeu 136
|||||
417 TTAGGCATCTCTAAGCCAAACATTTGATTTAGTCTAAAGGATTTAG 466
136 splyValIleThrValGlnLysThrValThrAlaIleLysThrIleVal 152
|||||
467 ATAAAGTTATAACGTACAAAAAAGGTAAGTCTGATTAACCATTTGTT 516
153 IleuAspSerLysValAspTyrArgGlyTyrGlnSerMetAspAsn 169
|||||
517 ATATTGACAGCAAGATGATATAGAGTTATCATCATGACACACTT 566
169 eLLeLysLysAsnThrProGlnGlyPheLysGlySerSerPheLysThy 186
|||||
567 TATTAATAAATAACACTCCACAGGTTTCAAGATGATCAAGTTTAAACTG 616
186 alGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSerGly 202
|||||
617 TAGAGTTAAACCGCAAGCAAGTGTCTTATTAATGAACCTTGGGT 666
203 SerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnIleVal 219
|||||
667 TCAACCGTTTGCACAAAGGTGACAACTTACTCATGAAATTTGGTCA 716
219 rArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValSerPro 236
|||||
717 GCGTTTTCACAGCTAGATCCAAATTTATGAAACCAAGTTTCACAG 766
236 lYThrAlaIleLeuThrValValProPheHisGlyPheGlyMetPhe 252
|||||
767 GCACGGCTATTTAAGTATGATGATGATGATGATGATGATGATGAT 816
253 ThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLeuThr 269
|||||
817 ACTACTTATGAGCTATCTAAGTGTGTTTTCGATGATGATGATGAT 866
269 sPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysCysSer 286
|||||
867 ATTTGACAGAGAGACTTTTAAACACCTGCACAGATTAACAATGTTCA 916
286 eValIleLeuValProThrLeuPheAlaIleLeuAsnArgSerGluLeu 302
|||||
917 GCGTATCTCTGACCGACTTGTGCAATCTTAATGAAAGTGAATTA 966
303 LeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGlyVal 319
|||||
967 CTGATTAATATGATTTATCAAAATTTAGTGAATTCATTCGCGGAGC 1016
319 aProLeuSerLysGluIleGlyValAlaValAlaArgArgPheAsnLeu 336
|||||
1017 ACCTTATCTAAGAAATGATGTAAGCTTCTCTGACCTTTTATTTAT 1066
336 roGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAlaIleLe 352
|||||
1067 CCGGTCTGTCAGAGCTATGTTTAAACAGAAACAACCTGCAATTTT 1116
353 lIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysValVal 369
|||||
1117 ATCACACCGGAAGGAGATGATTAACCAAGTGTCTGCAAGTGTGCC 1166
369 oLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGly 386
|||||
1167 ATTATTTAAGCAAAAGTTATGATCTTGAATACTAAAAAATCTTGGGC 1216
386 roAsnArgArgGlyGluValLysValLysGlyProMetLeuMetLysGly 402
|||||
1217 CGAACAGACGTGAGAAAGTTGTGTAAGGCTCCATCTATATAAAGT 1266
403 TyrValAspAsnProGluAlaThrArgGluIleLeuAspGluGluGly 419
|||||
```

```

1267 TATGTAGATATATCCAGAACACAGAGAAATCATATGATGAGAGCTTG 1316
419 PLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisPhePheI 436
1317 GTTGCACACAGAGATATTTGGTATTCAGATGAGAGAAAAAATCTTTTA 1366
436 LeValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGlnValPro 452
1367 TCGTGATCGCTTGAAGTCTTATATCAAAATACAAAGATATCAAGTACCA 1416
453 ProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePheAspAl 469
1417 CCGCTGTAATTAAGATCTGTTCTTTTGCACATCCAAATATTTTGTATGC 1466
469 aGlyValAlaGlyValProAspProIleAlaGlyGluLeuProGlyValAla 486
1467 CGGGGTGCGGGGCTTCACATCTATAGCTGTGAGACTTCGGGAGCTG 1516
486 aValValLeuLysLysGlyLysSerMetThrGluLysGluValMetAsp 502
1517 TTGTTGACTTGAAAAAGAAAAATCTATGACTGAAAAAGAAATGATGAT 1566
503 TyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyGlyValAla 519
1567 TACGTTGCTAGTCAAGTTTCAATGCAAAACGTTTGGCTGTGTGTCGCG 1616
519 gPheValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLysA 536
1617 TTTTGTGACAGAGTACTTAAGGTCTCAGTGAATTAATGACGGTAAG 1666
536 IaIleArgGluIleLeuLysLysProValAlaLysMet 548
1667 CAATTAAGAAATCTGAGAAACCAAGTGTCTAAGATG 1704

```

seq_name: /cgn2_6/plodata/2/lna/5A_COMB.seq:US-07-675-211-1

```

seq_documentation_block:
; Sequence 1, Application US/07675211
; Patent No. 5219737
; GENERAL INFORMATION:
; APPLICANT: KAJIYAMA, NAOKI
; TITLE OF INVENTION: MUTANT LUCIFERASE OF A FIREFLY, MUTANT
; TITLE OF INVENTION: LUCIFERASE GENES, NOVEL, RECOMBINANT DNAs CONTAINING THE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: N.Y.
; COUNTRY: U.S.A
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/675,211
; FILING DATE: 19910326
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MISHROCK, S. DESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7005-026-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1644 base pairs

```

```

; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Luciola cruciata
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1644
; US-07-675-211-1

```

```

alignment_scores:
  Quality: 2684.00      Length: 548
  Percent Similarity: 99.088      Gaps: 0
  Percent Identity: 93.431

```

alignment_block:

US-09-581-241-6 x US-07-675-211-1 ..

```

Align seg 1/1 to: US-07-675-211-1 from: 1 to: 1644

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17
1 ATGGAAACATGGAAGGATGCGTGAACACAAATTAACGCAATACATGG 50
17 eTyrProIleGluGluGlySerAlaGlyAlaGlnLeuArgLysTyrMetA 34
51 TTACCTATTCGAAAGGATGCGTGAACACAAATTAACGCAATACATGG 100
34 sParGThrAlaLysLeuGluAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 AGCGATATCGAAACACTTGCGCAATGCTTTACAAATCAGTACTGCT 150
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysLeuGly 67
151 GTGATATATCTTACCGCGAATCTTGAGAAATCATGTTGCTAGAGAA 200
67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgGlieAlaLeuCys 84
201 AGCTTGCAAAATTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 250
84 eGluAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
251 GTGAAACCTGGAAGATTTTATTCCTGATATACCGGACTGTTATA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgLeuVal 117
301 GGGTAGCTGTGCACCCACTAATGAGATTACACTTACGTAAGTGGT 350
117 HisSerLeuGlyIleSerLysProThrIleValPheSerSerLysG 134
351 TCACAGTTTAGGATCTCTAACCACCAATTTATTTAGTTCAAAAAAG 400
134 LysAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
401 GCTTAGATAAAGTATTAACAGTACAGAAACAGTACTATTAATAACC 450
151 IleValIleLeuAspSerLysValAspTyrArgLysTyrGlnSerMetA 167
451 ATTGTATACAGTACCAAAAGTATTAATGAGATATCAAGTCTGGA 500
167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheI 184
501 CACCTTTATAAAGAAACACTCCACAGGTTTCAAGCATCCAGTTTCA 550
184 yStrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
551 AAAGTGGAGAGTGAACCGTAAAGAACAGTGTCTTAATGAAGAACTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnI 217
601 TGGGTTCTACCGGTTTGCCAAAAGGCGTACACTTACTCACGAAATTAC 650

```

```

217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
|||||
651 AGTACTAGATTCTTCATGCTAGAGATCGATTATGTAACCAACTTT 700
234 ePProGlyThrAlaIleLeuThrValValProPheHisGlyPheGly 250
|||||
701 CACCAAGCACCGCTGTTTAACTGCTCCATTCATCATGCTTTGCT 750
251 MetPheThrThrLeuGlyTyrIleuThrCysGlyPheArgIleValMet 267
|||||
751 ATGTCACACTCTAGCGTATTATTAATTTGGTTCGTTGTTGTAAGTT 800
267 uThrIysPheAspGluGluThrPheLeuIysThrLeuGlnAspTyrLysC 284
|||||
801 AACAAATTCGATGAGAAACATTTTAAACTCTACAGATTAATAAT 850
284 ySSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
|||||
851 GTACAGTGTATCTTGTACGACCTGTTTGCATTCCTCAACAAAGT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
|||||
901 GAATTAACCTAATAAATACGATTGTCATTAATTAGTTGACATTGCATCTG 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheA 334
|||||
951 CGGACACCTTATCAAAAGAGTGTGACCTGTTGCTAGACGCTTA 1000
334 sNLeuProGlyValAlaArgGlnIlyTyrGlyLeuThrGluThrHisAla 350
|||||
1001 ATCTCCCGGCTGCTCTCAAGTTATGTTTACAGAAACCAATCTGCC 1050
351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
|||||
1051 ATTATTATTACACGAGAAAGACATTAACCAAGAGCTTGGAAAGT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
|||||
1101 CGTGCCGTTGTTTAAAGCAAAAGTTATGATCTTATACCAAAAATCTT 1150
384 euGlyProAsnArgArgLysGluValCysValLysGlyProMetLeuMet 400
|||||
1151 TAGGCTTAACGACGTGAGAGAGTTTGTAAAGACCTATGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGluG 417
|||||
1201 AAAGTTATGTAATAATCAACAGAACCAAAAGAACTTATGACGAGA 1250
417 uGlyTyrPheuHisThrGlyAspIleGlyTyrTyrAspGluGluLysH 434
|||||
1251 AGGTTGGCTGCACACGAGATATTGATATTATGATGAGAAACATTT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
|||||
1301 TCTTATATGTCGATCGTTGAAGCTTTAATCAATCAAAAGATACCAA 1350
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIleP 467
|||||
1351 GTACCACTGCCGAATTAAGAAATCGTTCTTTTGCACAACTCATATCTT 1400
467 eaSPAlaGlyValAlaGlyValProAspProIleAlaGlyLysLeuProG 484
|||||
1401 TGATGCTGGTGTGCGCGCTTCTGATCCGTGAGCTGCGACGCTTCAG 1450
484 lYAlaValValValLeuLysGlyLysSerMetThrGluLysGluVal 500
|||||
1451 GAGCGCTGTTGTTACTGAAAGCGGAAATAATATGACCGAAAGAACTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyG 517
|||||
1501 ATGATATTATGTCAGAGTCAAGTTTCAAAATGCAAAACGTTTACGTGTG 1550

```

```

517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
|||||
1551 TGTTCGTTTGTGATGAGAAATACCTAAAGGCTCTTACTGAGAAATTCAGC 1600
534 lYAlaValAlaArgGluIleLeuLysLysProValAlaLysMet 548
|||||
1601 GCAGAGCAATTAGAGAAATCTTAAAGAACCACTGCTTAAGATG 1644
seq_name: /cgn2_6/ptodata/2/lna/5A.COMB.seq:US-07-903-047-1
seq_documentation_block:
; Sequence 1, Application US/07903047
; Patent No. 5229285
; GENERAL INFORMATION:
; APPLICANT: Kajiyama, Naoki
; APPLICANT: Nakano, Eiichi
; TITLE OF INVENTION: Thermostable luciferase Of Firefly,
; TITLE OF INVENTION: Thermostable luciferase Gene Of Firefly, No. 5229285el Reco
; TITLE OF INVENTION: DNA, And Process For The Preparation Of Thermostable
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennile & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/903,047
; FILING DATE: 19920623
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7005-048
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1644 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA to mRNA
; US-07-903-047-1
alignment_scores:
Quality: 2684.00 Length: 548
Ratio: 4.943 Gaps: 0
Percent Similarity: 99.088 Percent Identity: 93.431
alignment_block:
US-09-581-241-6 x US-07-903-047-1 ..
Align seg 1/1 to: US-07-903-047-1 from: 1 to: 1644
1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProph 17
|||||
1 ATGCAAAACATGCAAAACGATGAAATATTTGTAGTTGACCTTAAACGTT 50
17 eTyrProIleGluGluGlySerAlaGlyAlaGlnLeuArgLysTyrMet 34
|||||
51 TTACCTATATGCAAGAGGATCTGCTGAGACACCAATTAACGCAATACATG 100

```



```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/076.042
FILING DATE: 15-JUN-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/675,211
FILING DATE: 26-MAR-1991
NAME: 435
ATTORNEY/AGENT INFORMATION:
NAME: MISROCK, S. LESLIE
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7005-026-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1644 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Lnc101a cruciata
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1644
US-08-076-042-1

```

```

alignment_scores:
Quality: 2684.00      Length: 548
Ratio: 4.943          Gaps: 0
Percent Similarity: 99.088      Percent Identity: 93.431

```

```

alignment_block:
US-09-581-241-6 x US-08-076-042-1 ..

```

```

Align seg 1/1 to: US-08-076-042-1 from: 1 to: 1644

```

```

1 MetGuaaMeGluAsnAspGluAsnIleValIYrGlyProGluProPh 17
1 ATGGAAACATGGAAGAAACGATGGAATTTGTTGACCTTAACCGCTT 50
17 eTyProlleGluGluGlySerAlaGlyAlaGlnLeuArgLysTyrMet 34
51 TTACCCCTTCGAGAGGATCTGCTGACACACATTAACGCAATACATGG 100
34 sPArGTYrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 AGCGATATGCAAAACCTTGCGCAATGCTTTACAAATGCAAGTACTGT 150
51 ValAspTYrThrAlaGluTyrLeuGluLysSerCysGlyLeuGlyG 67
151 GTTATATTCTTACGCCGAATACTTGAGAAATCATGTTGTTAGGAAA 200
67 uAlaLeuLysAsnTYrGlyLeuValAlaLysGlyArgIleAlaLeuCys 84
201 AGCTTTCGAAATATTATGGTTGGTTGGATGGAGCATTCGCTTATGCA 250
84 eArgLysGlyGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
251 GTGAAACCTGGAAGAAATTTTATCTCTTAATAGCCGACGCTTTATA 300
101 GlyValGlyValAlaProThrAsnGluLeuTyrThrLeuArgGluLeu 117
301 GGTGTAGGTGTGCAACCCACATGAGATTTACCTTACGTAACCTGCT 350
117 lHsSerLeuGlyLysSerLysProThrIleValPheSerSerLysLysG 134

```

```

351 TCACAGTTTAGGTATCTCTTAACCAACATTTAGTTCTTAACAAAG 400
134 lYeuAspLysValIleThrValGlnLysThrValPheAlaIleLysThr 150
401 GCTTACATAAAGTTATACAGTACAGAAACAGTAACACTACTATTAAACC 450
151 lIleValIleLeuAspSerLysValAspTYrArgLysTyrGlnSerMet 167
451 ATTGTTATCTAGATAGCAAAAGTTGATTCGAGATATCAATCAATGTCGA 500
167 pAsnPheIleLysAsnThrProGlnGlyPheLysGlySerSerPheI 184
501 CACCTTTATATAAAGAAACACTCCACAGGTTTTCAGATCCAGTTTCA 550
184 ySthrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
551 AAACGTGGAAGTTGACCGTAAAGAACAGTTGCTTATTAATGAACTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGlnAsn 217
601 TCGGTTCTACCGGTTTCCCAAAAGCGTACAACTTACGCAAAATAC 650
217 eValThrArgPheSerHisAlaArgAspProIleTYrGlyAsnGlnVal 234
651 AGTCACTAGATTTTCTCATGCTAGATCCGATTTATGTTAACCAAGTT 700
234 ePProGlyThrAlaIleLeuThrValAlaProPheHisIleGlyPheGly 250
701 CACCAAGCACCGCTGTTTAACTGTCGTCATTCATCATCATGTTTGGT 750
251 MetPheThrThrLeuGlyTYrLeuThrCysGlyPheArgIleValIle 267
751 ANGTTCACACTACTAGGATTTATTTATTTGTTTTCGTTGTTATGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTYrLysC 284
801 AACAAATTCGATGAAGAAACATTTTAAACCTACAAAGATTAAT 850
284 ySSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTTCAAGTTTATCTCTGACCGACCTGTTTGCATTTCTCAACAAAGT 900
301 GluLeuLeuAspLysTYrAspLeuSerAsnLeuValGluIleAlaSerG 317
901 GAATTCATCAATTAATACGATTTGTCAAATTTAGTTGATGATGATCG 950
317 yGlyAlaProLeuSerLysGluIleGlyGlnAlaValAlaArgArgPhe 334
951 CGGACACCTTTATCAAAAGAGTTGGAGCTGTGCTAGACGCTTTTA 1000
334 snLeuProGlyValArgGlnGlyTYrGlyLeuThrGluThrThrSerAla 350
1001 ATCTTCGCGGCTTGTGTAAGTTATGCTTTTACGAAACACATCTCC 1050
351 lIleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLys 367
1051 ATTATTTATACACAGAGAGAGATTAACACAGAGCTTCGGAAGAGT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThr 384
1101 CGTGCGCTGTTTAAAGCAAAAGTTATGATCTGATACCAAAAATCTT 1150
384 eGlyProAsnArgArgLysGluValCysValLysGlyProMetLeuMet 400
1151 TAGGTCCTAACAGAGCTGAGAGATTTGTTAAAGACCTATGCTTATG 1200
401 LysGlyTYrValAspAsnProGluAlaThrArgLysIleIleAspGlu 417
1201 AAAGTTATCTTAATTAATCCAGAACCAAAAGAACTTATGACGAAGA 1250
417 uGlyTyrLeuHisThrGlyAspIleGlyTYrTYrAspGlnLysHisP 434
1251 AGTTGGCTGCACACCGAGATTTGGATTTATGATGAGAAAGAAACATT 1300

```

```

434 hephelevalaspargleuylserleuileuyltyrlysglytyrln 450
1301 TCTTATTGTCGATCGTTGAACTTTATCAATAAAGGATACCAA 1350
451 ValProprolaagiuleugluservalleuileuinhisprosnleph 467
1351 GTACCCACTCCGAAATAGAACCGTTCTTTGCAACATCCATCTATT 1400
467 easpaalaglyvalalaglyvalproaspproilealaglyleuuprog 484
1401 TGATGCGGTGGTCGGCGCTTCCTGATCCTGTAGCGGCGACTTCAG 1450
484 lylalaValalvalleuylslysglylsermethrghlylsglyval 500
1451 GACCCGTTGTTGTTACTGGAAGCGGAAAAAATATGACCGAAAAAGTA 1500
501 Metasppyrvalalaserlnvalaserlnalalysargleuarglyl 517
1501 ATGGATTATGTCGACGTCAAGTTTCAAAATGCAAAACGTTTACGTGG 1550
517 yValargpheValaspcluValprolysglyleuthrghlylsglyleasp 534
1551 TGTTCGTTTGTGATGAAGTAAAGCTTACTGAAAAATGACG 1600
534 lylsAlaileargluileuylslysglylserprovalalalysmet 548
1601 GCACAGCAATTAGAGAAATCCTTAAAGAACGAGTGTGAAGATG 1644
seq_name: /cgn2_6/prodata/2/lna/5B_COMB.seq:us-08-757-046A-3
seq_documentation_block:
; Sequence 3, Application US/08757046A
; Patent No. 5876995
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; TITLE OF INVENTION: BIOLUMINESCENT ARTICLES OF MANUFACTURE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/757,046A
; FILING DATE: 11-25-96
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/597,274
; FILING DATE: 02-06-96
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6680-105B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1644 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:

```

```

; NAME/KEY: Coding Sequence
; LOCATION: 1..1644
; OTHER INFORMATION: Luciola cruciata Luciferase (Firefly)
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: PATENT NO.: 4,968,613
; US-08-757-046A-3

alignment_scores:
  Quality: 2684.00      Length: 548
  Ratio: 4.943          Gaps: 0
  Percent Similarity: 99.088      Percent Identity: 93.431

alignment_block:
US-09-581-241-6 x US-08-757-046A-3 ..

Align seg 1/1 to: US-08-757-046A-3 from: 1 to: 1644

1 Metgluasnmecgluasnaspcgluasnilevaltyrghlyprogluproph 17
1 ATGGAAACATGGAAGAAACATGAAATATTGTAGTTGACCTAAACCGTT 50
17 eTyrProilegluglylserAlaglyalaglnleuarglytyrmeta 34
51 TTACCTATCGAAGAGGATCTGCTGGAACACAAATTACGCAATACATG 100
34 sparqlyralalysleuglylalailealaphethrasnialeuthrghly 50
101 AGCGATATGCAAAACTTGGCGCATTTGCTTTTCAAAATCAGTACTGCT 150
51 Valaspyrthrtyralagluylleuylslysglylsercysleuglyl 67
151 GTTGATTATCTTACCGCGAATCTTGAGAAATCATGTTGCTAGGAA 200
67 ualaleuylsasntyrghlyleuvalalaspjargyleuileucys 84
201 AGCTTGCAGAAATATGATGTTGTTGATGGCAAAATGGCTTAGCA 250
84 ergluasncysglugluhpheleprovalleuAlaglyleupeulle 100
251 GTGAAACCTGAAAGATTTTATTCCGTATATACCGGAGCTGTATATA 300
101 GlyvalglyvalAlaprothrasngluiletyrThrleuargluenva 117
301 GGGTAGGTGTGACCCACTAATGAGATTACACTTACGTGTAACGTG 350
117 lHisserleuglylserlyspthrlevalpserSerlylsg 134
351 TCACAGTTTAGTATCTTAACCAACAAATTTGATTAGTTCTTAAAAAAG 400
134 lyleuaspplysvallethrValgluInlythrValalaleuylsrhr 150
401 GCTTAGTAAAGTATATACAGTACAGAAACAGTACTATTAAAC 450
151 lilevalleleuaspserylvalaspyrarglytyrlnsermetlas 167
451 ATTGTATACTAGATACCAAAAGTTGATTATCGAGGAATTAAGTCTGA 500
167 pasnPhelleuylslysanthrproglnglylphelylserSerphel 184
501 CACCTTATTAAGAAAGAACACTCCACAGGTTTCAAGCATCCAGTTTCA 550
184 ysrhrValgluValasnarglylsgluInvalalaleuilemetasnsr 200
551 AAACGTGGAAGTTGACCGTAAAGCAAGTGTCTTATTAAGCAACTCT 600
201 SerghlyserthrghlyleuprolysglyValgluInleuthrghlysn 217
601 TCGGTTCTTACCGGTTGGCCAAAGGCGTACAACTTACTACGAAATAC 650
217 eValThrArgPheSerHisAlaArgaspproiletyrghlyAsnlnValS 234
651 AGTCACTAGATTTTCTCATGCTAGATCGATTTATGTATACCAACTTT 700

```

```

234 expProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
|||||
701 CACGAGCAGCAGCGCTGTTTAACGTGCTTCATTCATCATGTTTGGT 750
|||||
251 MetPheThrThrLeuGlyLeuThrCysGlyPheArgIleValMetIle 267
|||||
751 ATGTTTACTACTCTAGGGTAATTAATTTGTTGTTTCTGTTGTAATGTT 800
|||||
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
|||||
801 AACAAATTCGATGAGAGAAACATTTTAAAACTTACACAGATTATTAAT 850
|||||
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
|||||
851 GTACAGAGTGTATTTCTTTGACCGACCTGTTTGCATTTCTCACCAAAAG 900
|||||
301 GluLeuLeuAspLysTyrAspLysSerAsnLeuValGluIleAlaSerG 317
|||||
901 GAATTTACTCAATTAATACGATTTGTCAAATTTAGTTGAGATTGTCATCTG 950
|||||
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
|||||
951 CCGAGACACTTATTCAAAAGAGTTGTAAGCTGTTGCTAGACGCTTTA 1000
|||||
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrSerAla 350
|||||
1001 ATCTTCCGCGTGTGTCAGAGTTAGTTTAAACAGAAACAACATCTGCC 1050
|||||
351 IleIleIleThrProGluGluValAspAspLysProGlyAlaSerGlyVal 367
|||||
1051 ATTATTATTACACCAAGAGACATTAACCAAGAGCTTGGGAAAAGT 1100
|||||
367 ValProLeuPheLysAlaLysValIleAspLeuAspThrLysThrL 384
|||||
1101 CGTGCGCTGTTTAAAGCAAAAGTATTGATCTGTGATACCAAAAATCTT 1150
|||||
384 euGlyProAsnArgArgGlyGluValLysValLysGlyProMetLeuMet 400
|||||
1151 TAGGTCCTTAACGACGTGAGAGAGTTTGGTTAAAGAGACCTTAGCTTAG 1200
|||||
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGluG 417
|||||
1201 AAAGGTATGTAATAATCCAGAACCAACAAAGAACTTATTGACAGAA 1250
|||||
417 uGlyTyrPheLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLys 434
|||||
1251 AGGTTGGCTGCACACCGAGATATTGATGATGATGAGAAAAACATTT 1300
|||||
434 hepPheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
|||||
1301 TCTTTATTGTGATCGTTTGAAGTCTTTAATCAATTAACAAAGGATACCA 1350
|||||
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIleP 467
|||||
1351 GTACCACCTCGCCGAAATAGATCCGTTCTTTGCAACATCATCATCTT 1400
|||||
467 eaSPAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
|||||
1401 TGAATGCTGGGTGTGCGCGGCTTCCGATCTGTGCTGCGAGCTTCCAG 1450
|||||
484 lYalAlaValAlaValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
|||||
1451 GAGCGGTTGTTGTTACTGAAAGCGGAAAAAATATGACCGGAAAAAGAACTA 1500
|||||
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyG 517
|||||
1501 ATGATATTATGTTGCAGTCAGATTCAAAATCAAAAGCTTTACGCGTGG 1550
|||||
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
|||||
1551 TGTTGTTTTGTGGATGAGTACTAAAGTCTTACTGGAAGAAAAATTGACG 1600
|||||

```

```

534 lYlSaAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
|||||
1601 GCAGAGCAATTAGAGAAATCTTAAAGAAACCACTTGCTAAGATG 1644
|||||
seq_name: /cgn2-6/plodata/2/ina/6A_COMB.seq:US-09-447-208-3
seq_documentation_block:
; Sequence 3, Application US/09447208
; Patent No. 6113886
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; TITLE OF INVENTION: BIOLUMINESCENT ARTICLES OF MANUFACTURE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McCauliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/447,208
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 0909/135,988
; FILING DATE: 08-17-98
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/757,046
; FILING DATE: 11-25-96
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/597,274
; FILING DATE: 02-06-96
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24727-105C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-450-8400
; TELEFAX: 619-450-8499
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1644 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...1644
; OTHER INFORMATION: Luciola Cruciata Luciferase (firefly)
; PUBLICATION INFORMATION:
US-09-447-208-3
alignment_scores:
Quality: 2684.00 Length: 548
Ratio: 4.943 Gaps: 0
Percent Similarity: 99.088 Percent Identity: 93.431
alignment_block:
US-09-581-241-6 x US-09-447-208-3 ..
Align seg 1/1 to: US-09-447-208-3 from: 1 to: 1644

```

```

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPhe 17
1 ATGGAAGAAACATGAGAAACATGAGAAATATGTAGTGGACCTAAACCGTT 50
17 eTyrProIleGluGlySerIleGlyIleValIleuArgLysTyrMetA 34
51 TTACCCCTATGCAAGAGGATCTCTGGACACATTAACGCAATTAACATGG 100
34 sPArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 AGCATATGCAAAACTTGGCGCAATGCTTTACAATGCAAGTACGTTACGT 150
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGly 67
151 GTGATATATCTTACGCCGATCTGGAGAAATCAATGTTGTCTAGAGAA 200
67 uAlaLeuLysAsnTyrGlyLeuValIleAlaPheArgIleAlaLeuCys 84
201 AGCTTTCGAAATATGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 250
84 eArgLysAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
251 GTGAACCTGGAAGAAATTTTATCTCTGTAATAGCCGACGCTGTATA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
301 GGTGATGCTGTGGACCCACCAATGAGATTTACCTTACGTAAGTGGT 350
117 HisSerLeuGlyIleSerLysProThrIleValPheSerSerLysLys 134
351 TCACAGTTAGTATCTCTTAACCAACAATGTTAGTCTTAAAG 400
134 LysLeuAspLysValIleThrValGluLysThrValIleAlaIleLysThr 150
401 GCTTATGATTAAGTATATACAGTACAGAAACAGTAACCTACTATTAAACC 450
151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAs 167
451 ATTGTATTAATAGTATGCAAGGATGATTAATGAGATATCAATGCTGGA 500
167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerPheIle 184
501 CACCTTTATTAAGAAACACCTCCACGAGTTTCAACGATCCAGTTTCA 550
184 YstHValGluValAsnArgLysGluIleValAlaIleLeuIleMetAsnSer 200
551 AAACGTGTGAAGTGTACCGTAAAGCAAGTGTCTTATTAATGAACCT 600
201 SerGlySerThrGlyLeuProLysGlyValGluLeuThrHisGluAsnIle 217
601 TCGGGTTCACCGGTTTGCAGAAAGCGGTACACACTTACACAGAAATAC 650
217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
651 AGTCACTAGATTTTCTCATGCTAGAGATCCGATTTATGTATGTAACCAAGTT 700
234 ePrGlyThrAlaIleLeuThrValIleProPheHisGlyPheGly 250
701 CACCAAGCACCGCGTGTATTAACGTCGCTTCATTCATCATGTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetIle 267
751 ATGTTACACTACCTTACGATTAATTTATTTGTGTTTCTGTTGTTATCT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
801 AACAAATTCGATGAAGAAACATTTTAAACCTCAACAGATTAATAAT 850
284 YSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTACAAAGTGTATCTTGTATGACGACCTGTTGTCATTCCAACAAAG 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317

```

```

901 GAATTAATCAATAATTAATGATTTGTCAATATTAGTGGATGATCG 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheA 334
951 CGAGCACCTTATCAAAAGAAAGTGTGTAACCTGCTGACGCTTTA 1000
334 sPLeuProGlyValAlaArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
1001 ATCTTCCGCGGTGTCTCAAGTTATGTTTAAAGAAACAAACATCTGCC 1050
351 IleIleIleThrProGluLysAspLysProGlyAlaSerGlyLysVal 367
1051 ATTATATTAACACAGAAAGAGACGTAACACAGAGCTTCTGAAAAAGT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
1101 CGTCCGCTGTTTAAAGCAAAAGTATGATCTGTATACCAAAATATCTT 1150
384 eGluProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
1151 TAGTCTTAACAGACGTGGAGAAATTTGTTAAAGACCTATGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGlu 417
1201 AAAGTTATGTAATATATCAAGAACCAACAAAGAACTTATGACGAGA 1250
417 uGlyTyrPheHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
1251 AGGTGTGCTCACCGGAGATATGATATTTATGATGAAGAAAAACATT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
1301 TCTTATATGTCGATCGTTGAAGCTTTATCAATCAAAAGATACCAA 1350
451 ValProProAlaGluLeuGluSerValLeuGluHisProAsnIlePh 467
1351 GTACACCTCCGCAATTAAGATCCGTTCTTTCACATCTCATCTTCTT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyLeuProG 484
1401 TGATGCGGTGTGTCGCGGCTCTCTGATCTGTAGCTGGAGACTTCAG 1450
484 lYalaValAlaValLeuLysLysGlyLysSerMetThrGlyLysGluVal 500
1451 GAGCGTGTGTGTACTGCAAGCGGAAAAATATGACCGCAAAAAAGATA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
1501 ATGATTAATGTTGCAAGTCAAGTTTCAAAATGCAAAACGTTTACGTGG 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
1551 TGTTCGTTTGTGATGAGTAAAGTCTTACTAGCAAAATATGACG 1600
534 lLysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
1601 GCAGAGCAATTAGAAATCTTAAAGAACCAAGTTGCTAAGATG 1644

```

seq_name: /cgn2_6/plodata/2/lna/6A_COMB.seq:US-09-135-988-3

seq_documentation_block:

Sequence 3, Application US/09135988

Patent No. 6152358

GENERAL INFORMATION:

APPLICANT: Bryan, Bruce

TITLE OF INVENTION: BIOluminescent ARTICLES OF MANUFACTURE

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Heller Ehrman White & McCaulliffe

STREET: 4250 Executive Square, 7th Floor

CITY: La Jolla

STATE: CA

```

? COUNTRY: USA
? ZIP: 92037
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FastSeq Version 1.5
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/135,988
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/757,046
? FILING DATE: 11-25-96
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/597,274
? FILING DATE: 02-06-96
? ATTORNEY/AGENT INFORMATION:
? NAME: Seidman, Stephanie L
? REGISTRATION NUMBER: 33,779
? REFERENCE/DOCKET NUMBER: 24727-105C
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 619-450-8400
? TELEFAX: 619-450-8499
? TELEX:
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1644 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? FEATURE:
? NAME/KEY: Coding Sequence
? LOCATION: 1..1644
? OTHER INFORMATION: Luciola Cruciata Luciferase (Firefly)
? PUBLICATION INFORMATION:
? US-09-135-988-3

alignment_scores:
  Quality: 2684.00      Length: 548
  Ratio: 4.943          Gaps: 0
  Percent Similarity: 99.088      Percent Identity: 93.431

alignment_block:
  US-09-581-241-6 x US-09-135-988-3 ..

Align seg 1/1 to: US-09-135-988-3 from: 1 to: 1644

1 MetGluSnMetGluAsnAspGluAsnIleValTyrGlyProGluProph 17
  |||||||
1 ATGAAAAACATGAAAAACGATGAAATATGTAGTGGACCTAAACCGTT 50
  |||||||
17 eTyrProIleGluGlySerAlaGlyAlaGluLeuArgLysTyrMet 34
  |||||||
51 TTACCTATGCAAGAGGATCTGCTGCAACAACATTACGCAATACACTG 100
  |||||||
34 sPArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThgLy 50
  |||||||
101 AGCGATATGCAAAACCTTGCGCAATTCCTTTACAAATGCAAGTTACTG 150
  |||||||
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysLeuGlyG 67
  |||||||
151 GTGATATATCTTAACGCCGAATCTTGAGAAATCATGTTGCTAGGAAA 200
  |||||||
67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
  |||||||
201 AGCTTTGCAAAATTAAGTTGGTTGGTTGATGGCAGAAATGCGTATGCA 250
  |||||||
84 eGluLysCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
  |||||||

```

```

251 GTGAAACTGTGAAGATTTTTCCTGTAATAGCCGACGCTTTATA 300
  |||||||
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVa 117
  |||||||
301 GGCTAGAGGTTCACCCACTAATGAGATTACACTTTAGCTGACAGCTGT 350
  |||||||
117 LHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysG 134
  |||||||
351 TCACAGTTTAGGTATCTTAACCAACCAAAATGTAATTTAGTTCTAAAG 400
  |||||||
134 LLeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
  |||||||
401 GCTTAGATAAAGTTATTAACAGTACAGAAACAGTAAGTATTTAAACC 450
  |||||||
451 ATTGTATACATAGATAGCAAAAGTTGATTATCGAGATATCAATGCTGCA 500
  |||||||
167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPhe 184
  |||||||
501 CACCTTTATAAAAAGAAACACTCCACCAGGTTTCAAGCAAGTTTCA 550
  |||||||
184 YThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
  |||||||
551 AAAGTGTGAAGTTGACCGTAAAGAACAAAGTCTCTTATTAAGAACTCT 600
  |||||||
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsn 217
  |||||||
601 TCGGTTCTACCGGTTTGGCCAAAGGCGTAACTACACCAAAATAC 650
  |||||||
217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnVal 234
  |||||||
651 AGTACTAGATTTTTCATGCTAGATCGATTCGATTATGTAACCAAGTT 700
  |||||||
234 eProGlyThrAlaIleLeuThrValValProPheHisGlyPheGly 250
  |||||||
701 CACCAAGCACCGCTGTTTAACTGCTTCATCCATCATCATGATGTTGGT 750
  |||||||
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetI 267
  |||||||
751 ATGTCACACTCTAGGATATTTAATTTGGTTTCGTTGTAATGTT 800
  |||||||
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
  |||||||
801 AACAAATTCGATGAAGAAACATTTTAAAACTACAAAGATTATTAAT 850
  |||||||
284 YSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
  |||||||
851 GTACAGTGTATTTCTGTACCGACCTGTTGCAATTCCTCAACAAAGT 900
  |||||||
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
  |||||||
901 GAATTCCTCAATAAATACGATTTGTCAAATTTAGTTGAGATTGCATCTG 950
  |||||||
317 YGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
  |||||||
951 CGAGACCTTATATCAAAAGAGTTGGTAACCTGTTCTGAGACCTTTA 1000
  |||||||
334 sLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThSerVal 350
  |||||||
1001 ATCTCCCGGCTTCTGTCAGGTTATGGTTTACAGAAACACATCTCC 1050
  |||||||
351 IleIleIleThrProGlnGlyAspAspLysProGlyAlaSerGlyLysVa 367
  |||||||
1051 ATTATTATTACACGAGAGAGACGATTAACACAGAGCTTGGAAAAAGT 1100
  |||||||
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysTyrTr 384
  |||||||
1101 CGTCCGCTGTTTAAAGCAAAAGTTATTGATCTGATACCAAAAAATCTT 1150
  |||||||
384 eGluProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
  |||||||
1151 TAGTCTTAACAGACGTGAGAAAGTTTGTGTTAAAGACCATCTATG 1200
  |||||||

```



```

267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
801 AACAAATTCGATGAGAAACATTTTAAAACTCATACAGATATAAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuSnaArgSer 300
851 GTACAGAGTGTATCTGTACCGACCTTGTTCGAATCTCAACAAACAT 900
301 GLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
901 GAATTAATCTCAATATACGATTTGTCAAATTTAGTTGACATTCATCTG 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
951 CGAGAGACCTTATCAAAAGAGTTGTGAGCTGTTGCTAGACGCTTTA 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
1001 ATCTTCCCGGTGTGTCAGGTTATGTTTAAACGAAACCAATCTGCC 1050
351 IleIleIleThrProGlnGlyAspAspLysProGlyAlaSerGlyLysVa 367
1051 ATTATATTATACACGAGAGACCATTAACAGAGCTTCTGAAAACT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysThrL 384
1101 CGTGGCGTGTTTAAAGCAAAAGTTATGATCTGTATACCAAAAATCTT 1150
384 euGlyProAsnArgArgGlyGluValGlyValGlyGlyProMetLeuMet 400
1151 TACGTCCTACAGACGTGAGAGAGTTGTGTTAAAGACCTATGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleLeuAspGluG 417
1201 AAAGGTATATGTAATATATCCAGACCAACAAAGAACTTATGACAGAGA 1250
417 uGlyThrPheuHisThrGlyAspIleGlyTyrTyrAspGluGluLysH 434
1251 AGGTTGGCGGCACCGGAGATATGATATGATGATGATAAAGAAACAT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
1301 TCTTATATGTGATGCTTTGAAGTCTTTATACAAATTAACAAAGATACCA 1350
451 ValProProAlaGluLeuGlnSerValLeuLeuGlnHisProAsnIlePh 467
1351 GTACACCTGCCGAATTAGAATCCGTTCTTTCGCAACATCCATCTATCTT 1400
467 eaSpAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
1401 TGATGCTGTGTTGCCGCGCTTCTGATCTGTTAGCTGGCGAGCTTCCAG 1450
484 lAlaValAlaValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
1451 GAGCGCTGTGTTGACTGGAAGCGGAAAAAATATGACCGCAAAAAAGAGTA 1500
501 MetAspTyrValAlaIleSerGlnValSerAsnAlaLysArgLeuArgGlyG 517
1501 ATGGATTTATGTTGCAGATCAAGTTTCAAAATGCAAAAGCTTACGTTG 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleSpG 534
1551 TGTTCGTTTGTGATGACATCACTAAAGGCTTACTGAGAAAAATTAACG 1600
534 lYlyAlaAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
1601 GCAGAGCAATTAGAGAAATCTCTTAAGAACACAGTTGCTAAGATG 1644
seq_name: /cgn2_6/ptodata/2/1na/6B_COMB.seq:US-08-597-274A-3
seq_documentation_block:
; Sequence 3, Application US/08597274A

```

```

; Patent No. 6247995
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; TITLE OF INVENTION: BIOLUMINESCENT NOVELTY ITEMS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/597,274A
; FILING DATE: 02/06/96
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6680-105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1644 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1..1644
; OTHER INFORMATION: Luciola Cruciata Luciferase (Firefly)
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: 4,968,613
; US-08-597-274A-3

alignment_scores:
    Quality: 2684.00      Length: 548
    Ratio: 4.943          Gaps: 0
    Percent Similarity: 99.088      Percent Identity: 93.431

alignment_block:
US-09-581-241-6 x US-08-597-274A-3 ..
Align seg 1/1 to: US-08-597-274A-3 from: 1 to: 1644

1 MetGlnAsnMetGlnLysAspGlnLysAsnIleValTyrGlyProGluProPh 17
|||||
1 ATGGAACAATGGAAGAACGATGAAATATTTAGTTGACCTAAACCGTT 50
17 eTyrProIleGluGluGlySerAlaGlyAlaGlnLeuArgLysTyrMetA 34
|||||
51 TTACCTATCGAAGAGGATCTGTGGAAACAAATTACCGCAATACATG 100
34 sPArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
|||||
101 AGCGATATCAAAACCTGGCGCAATGCTTTTACAAATGCAATGCTACTG 150
51 ValAspTyrThrTyrAlaGluThrLeuGlnLysSerGlyLysGluGlyG 67
|||||
151 GTTGATTTATCTTACGCCGCAATCTTGGAGAAATCATGTTGTAGGAAA 200

```

```

67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCysS 84
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 AGCTTTCAGAAATTAATGTTGGTTGGTTGATGGCAGAAATTCGTTAAGCA 250
84 ergLysAsnGlyGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
251 GTGAAACCTGTGAAGAAATTTTATTCCTTAATAGCCGACGCTTTATA 300
101 GlyAlaGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
301 GGTTAGAGTGTGGCACCACCAATGACATTTACACTTTACGTGAACGTGCT 350
117 LHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysLysG 134
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
351 TCACAGTTAGTATCTCTAAACCAACAATTTGATTTGTTCAAAAAAG 400
134 LysLeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
401 GCTTAGATAAAGTTATACAGCTACAGAAACAGTAACCTACTATTAAAAAC 450
151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAs 167
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
451 ATGTGTTACTAGATACAGAAAGTTGATTATCGAGATATCAATGCTCGA 500
167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheI 184
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
501 CACCTTTATTAAGAAACACTCCACAGAGTTTCAAGCATCCAGTTTCA 550
184 YStrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
551 AAACGTGAGAGTTGACCGTTAAACAACAAGTTGCTTATATGAACTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnI 217
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
601 TCGGCTTCTACCGGTTGGCAAAAGGCGTACACTACTCAGCAAAATAC 650
217 eValThrPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
651 AGTCACAGATTTTCTCATGCTAGAGATCCGATTTATGTTAACCAGATT 700
234 erProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
701 CACACGACGACCGGTGTTTAACTGCTTCATTCACATCATGCTTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
751 ATGTTCACTACTCTAGGCTATTTAATTTGTTGCTTTCGTTGTAATGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
801 AACCAAAATTCGATGAGAAACATTTTAAAACTCTACAGATTTATTAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
851 GTACAAGGTATTTCTGTGACGACCTTGTTGCAATTCCTCAACAAAGT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
901 GAATTACTCAATAATACGATTTGCAAAATTTAGTTGAGATTGCATCTGG 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheA 334
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
951 CGGAGCACCTTTATCAAAAAGAGTTGGTGAAAGCTTGTCTAGACGCTTTA 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1001 ATCTTCCCGGTGTCGTCAAAGTTATGTTTAAAGAAAAACACATCTGCG 1050
351 IleIleIleThrProGlnGlyAspAspLysProGlyAlaSerGlyLysVa 367
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1051 ATTATATTATACACAGAAAGAGACGATTAACAGAGAGCTTCTGAAAAAGT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1101 CGTGCCGTTGTTTAAAGCAAAAGTTATTCATCTTGATACCAAAAAATCTT 1150
384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1151 TAGGTCTTAACAGACGTCGAGAAAGTTTGTGTTAAAGACCTATGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGluGl 417
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1201 AAAGTTATGTAATTAATCCAGAAAGCAACAAAGAACTTATTGAGAGAA 1250
417 uGlyTyrLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1251 AGGTTGGCTGCACACCGGAGATATGATATGATGAGAAAAACAATT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1301 TCTTTATGTCGATCGTTGAAAGTCTTTATCAAAATACAAAGGATACCA 1350
451 ValProProAlaGluLeuGlnSerValLeuLeuGlnHisProAsnIlePh 467
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1351 GTACCACTGCGCAATTAATGATCCGTTCTTTGCAAAATCAATCTATCTT 1400
467 eaSPAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1401 TGATGCTGGTGTGCCGGGCTTCCTGATCCTGAGCTGCGAGCTTCAG 1450
484 lValAlaValValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1451 GACCCGTTGTTGTAAGGAAAGGAAAAAATTTGACCGAAAAAGAAATTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgLysG 517
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1501 ATGATTTATGTTCAAGTCAAGTTCAAAATGCAAAAGCTTTAGTGTGTGG 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1551 TGTTGCTGTTGTAAGTAAGTACCTTAAAGGTCTTACTGGAAAAATATGAG 1600
534 lLysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1601 GCAGAGCAATTAAGAAATCTTAAAGAAACAGTTGCTAAGATG 1644

```


84 erGluAsnCysGluuphepHeleProValIleuAlaGlyLeuPheHele 100
 251 GTGAAAGCTGTAAGAAATCTTTATTCCTGATATAGCCGGTTTATTATTA 300
 101 GlyValGlyValAlaProThrAsnGluIleGlyThrLeuArgGluLeuVa 117
 301 GGTGTCGGGTGGCTCCAACTAATAGATTTACACTTACCTGATGAAATGGT 350
 117 HisSerLeuGlyIleSerLysProThrIleValPheSerSerLysLysG 134
 351 TCACAGTTTAGGCATCTCTAAGCCACAAATGTTATTTAGTTCTTAAAAAG 400
 134 IlyLeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
 401 GATTAGATAAAGTTATATACCTACAAAAAACGGTAACGTATATAAAAAC 450
 151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAs 167
 451 ATTGTTATATTTGACACGCAAGGTGATTTATAGAGTTATCATCATGCA 500
 167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheL 184
 501 CAACCTTATTTAAAAAACAACCTCCACAAGGTTTCAAGGATCAAGTTTAA 550
 184 YSThrValGluValAsnArgLysGluGlnValAlaIleuIleMetAsnSer 200
 551 AAACGTGTAAGTTATACCGCAAGACAAAGTGTCTTATATATGCAACTCT 600
 201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLe 217
 601 TCGGGTTCAACCGGTTGCCAAAAAGTGTGCAACTTACATCATGAATTT 650
 217 UValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
 651 GGTACTAGATTTTCTCAGCTGAGATCCAAATTTATGAGAAACAGTTT 700
 234 erProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
 701 CACAGGACGCGGTATTTTAACGTAGTACCATTCCATCATGTTTGGT 750
 251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetIe 267
 751 ATGTTTACTTACTTGGCTTACTTACCTTGTGTTTCTGATTTGTCATGTT 800
 267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
 801 AAGCAATTTGACGAAGACACTTTTAAACACACTGCAGATTTACAAAT 850
 284 YSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
 851 GTTCAGACGTTATTTCTGTACGACCTTTGTTGCAATTTCTTATACAACT 900
 301 GlnLeuLeuAspLysTyrAsnLeuSerAsnLeuValGluIleAlaSerG 317
 901 GAATTTCTGATTAATATGATTTATCAAAATTTAGTTGAATTTGCACTCTG 950
 317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheA 334
 951 CGGAGACCTTTATCTTAAGAATTTGTAACCTGTTGTTAGACGTTTAA 1000
 334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
 1001 ATTTACCGGGGTTCCTCAAGGCTATGTTTAAACAGAAACAACCTCTGCA 1050
 351 IleIleIleThrProGluGlyAspAspLysProGlyValIleSerGlyLysVa 367
 1051 ATTATTTATACACCGGAAGCGCATATTAACCAAGGTGCTCTGSCAAAGT 1100
 367 IValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
 1101 TGTGCATATTTTAAAGCAAAAGTTATCGATCTGATCTAATAAAACTT 1150
 384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400

1151 TGGGCCCAACACACGCGGAGAGTTTGTGTAAAGGTCCTATGCTTAA 1200
 401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGluL 417
 1201 AAAGGTATGTACATATATCCAGAACACACAGAAATATGATGATGAAGA 1250
 417 uGlyTyrPheHisThrGlyAspIleGlyTyrTyrAspGluLysHisP 434
 1251 AGGTTGGTTGACACACAGAGATTTGAGTATGATGATGAAGAAAAACAT 1300
 434 hePheIleValAspArgLeuLysSerLeuIleLysTyrGlyGlyGln 450
 1301 TCTTTATCGTGATTCGTTTGAGAGCTTTTATATCAAAATACAAAGATATCA 1350
 451 ValProProAlaGluLeuGlnSerValLeuLeuGlnHisProAsnIleP 467
 1351 GTACACACCTGTGAATTTAGATCTGTTCTTTGCAACATCCAAATATTTT 1400
 467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
 1401 TGTATCGCGCGGTGCTGCGGTTCCAGATCCATATACCTGATGAGCTCCG 1450
 484 IValAlaValAlaLeuLysLysGlyLysSerMetThrGluLysGluVal 500
 1451 GAGCTGTGTTGTTACTTATAGAAAGAAAAATCATGACTGAAAGAAAGTA 1500
 501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
 1501 ATGAGATTACGTTGCTAGTCAAAATTTCAAAATGCAAAACCTTTGCGGTGG 1550
 517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
 1551 TGTCCGTTTGTGGACGAAGTACCTAAAGTCTCATGTTAAATTTGACG 1600
 534 IlyLysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
 1601 GTAAAGCAATTAGAGAAATACTGAAAGAAACAGTTGTAAAGATG 1644

seq_name: gb_pat:E36480

seq_documentation_block: E36480 1644 bp DNA PAT 07-FEB-2001

LOCUS E36480 1644 bp DNA PAT 07-FEB-2001

DEFINITION Luciferase and method for assaying intracellular ATP by using the same.

ACCESSION E36480

VERSION E36480.1 GI:13022687

KEYWORDS JP 1999239493-A/4.

SOURCE Luciola lateralis.

ORGANISM Luciola lateralis

REFERENCE Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga; Elateriformia; Cantharoidae; Lampyridae; Luciola.

AUTHORS Noriaki H.S.M.M.

TITLE Luciferase and method for assaying intracellular ATP by using the

JOURNAL Patent: JP 1999239493-A 4 07-SEP-1999;

COMMENT KIRKMAN CORP

OS Luciola lateralis

PN JP 1999239493-A/4

PD 07-SEP-1999

PF 21-DEC-1998 JP 1998363108

PR

PI NORIAKI HATTORI, SEIJI MURAKAMI

PC C12N15/09, C12N9/02, C12Q1/66, C12N9/02, C12R1.19, C12N15/00 CC

FEATURES

source

FT Key Location/Qualifiers

FT CDS Location/Qualifiers

1..1644

/db_xref="taxon:7052"

BASE COUNT 529 a 262 c 348 g 505 t

ORIGIN

alignment_scores:
 Quality: 2821.00 Length: 548
 Ratio: 5.148 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.818

alignment_block:
 US-09-581-241-4 x E36480 ..

Align seg 1/1 to: E36480 from: 1 to: 1644

```

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17
|||||
1 ATGGAAGAAATGAGACGATGAAATATGTTGTTGCTGCTGAACCAATT 50
17 eTyrProIleGluGluGlySerAlaGlyAlaGlnLeuArgLysTyrMetA 34
|||||
51 TTAACCTATTTGAAGAGGATCTGCTGAGACACAAATTGCGCAAGTATATGG 100
34 sPArGTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
|||||
101 ATCGATATCGAAGAACTGGAGCAATTGCTTTACTAAGCACTTACCGGT 150
51 ValAspTyrThrTyrAlaGluTyrLeuGlyLysSerCysCysLeuGly 67
|||||
151 GTCGATTATACGTACGCCGAATCTTAGAAAAATCATGCTCTCTAGGAGA 200
67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
|||||
201 GGCTTTAAAGAAATTATGTTGGTTGTTGATGAGACAAATTCGCTTATGCA 250
84 eGluAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
|||||
251 GTGAAAACGTGAAAGATTCTTATCTGTATTAACCGGTTTATTATA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
|||||
301 GGCTGGGCTGGCTCCAACTAATGAGATTTCACCTCTACGTGATTTGGT 350
117 lHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysLysG 134
|||||
351 TCACAGTTTGGCATCTCTAAGCCAAATTTGATTAGTTTAAAAAAG 400
134 LysLeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
|||||
401 GATTAGATAAAGTTATTAAGTACAAAAAACGGTAACCTGCTATTAAAAAC 450
151 IleValIleLeuAspSerLysValAspTyrArgIleTyrGlnSerMetAs 167
|||||
451 ATGTATTATTGGACGACCAAGTGAGTTATAGAGTTATCAATCCATGGA 500
167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheL 184
|||||
501 CAACTTATTAAAAAACAACCTCACAAAGGTTTCAAAAGCATCAAGTTTAA 550
184 YsThrValGluValAsnArgLysGlnValAlaLeuIleMetLysSer 200
|||||
551 AAACGTAGAGATTAAACCGCAAGAACAAAGTCTCTTATTAATGAACCTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLe 217
|||||
601 TCGGGTTTCACACCGGTTGCCAAAAAGTGTCAACTTACTCATGAATAATAT 650
217 uValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
|||||
651 CGTCACTAGATTCTTCACGCTAGATCCAAATTTATGGAACCAAGTTT 700
234 eTyrProGlyThrAlaIleLeuThrValValProPheHisLysGlyPheGly 250
|||||
701 CACGAGGACGGCTATTATTACTAGTACCATTCATCATGTTTGGT 750

```

```

251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
|||||
751 ATGTTTACTACTTAGGCTATCTTAACCTTGTTGTTTCTGATTTCATGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
|||||
801 AACGAATTTGACGAAGACACTTTTAAAAACACGCAAGATTACAAAT 850
284 YsSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
|||||
851 GTTCACGCTTATCTTCTTACCGCACTTGTTCCTCAATTTCTTAATAGAGT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGlnIleAlaSerG 317
|||||
901 GAATTTACTCGATTAATATGATTTATCAAAATTTAGTTCAAAATTCATCTGG 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheA 334
|||||
951 CGAGACACCTTATCTTAAGAAATTTGTTGAAGCTGTGCTAGACGTTTAA 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
|||||
1001 ATTTACCGGCTGCTGTCACAGGCTATGTTTACAGAAACCACTCTGCA 1050
351 IleIleIleThrProGluGluLysAspLysProGlyAlaSerGlyLysVa 367
|||||
1051 ATTATTATCACACCGGAAGGCGATGATAAACCAAGTGCTTCTGCAAAAGT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
|||||
1101 TGTGCCATTATTAAAGCAAAAGTTATCGATCTTGATCACTAAAAAACTT 1150
384 euGlyProAsnArgArgGlyGlyValLysValLysGlyProMetLeuMet 400
|||||
1151 TGGGCCCCACAGACGCTGAGCAAGTTTGTGTAAGGCTCATATCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGlu 417
|||||
1201 AAAGTTATGTGATTAATCCAGAACCAAGAGAAATCATGATGATGAGA 1250
417 uGlyThrLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
|||||
1251 AGCTTGGTTCACACAGAGATTTGGGTATTTACATATACAGAAAAACCTT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
|||||
1301 TCTTTATCGTGAATGCTTTGAAGTCTTTAATCAATACAAAGATATCAA 1350
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
|||||
1351 GTACCCACTGCTGAATTAGAATCTGTTTGGCAACATCCAAATATTTT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
|||||
1401 TGTATCCGCGCTTGGCTGGGCTTCCAGATCTTACTGTTGAGCTTCGGG 1450
484 lYalAlaValValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
|||||
1451 GAGCTGTGTTGTACTTAGAAAGAAATCTATGACGTAAGAAAAAGATA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyG 517
|||||
1501 ATGGATTACGTTGCTAGTCAAGTTTCAAAATGCAAAACGTTTCCGTGGTGG 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
|||||
1551 TGTCCGTTTTTGGAGAGAGTACCTTAAGTCTCACTGTTAAATTTGACG 1600
534 lLysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
|||||
1601 GTAAACCAATTAGAGAAATACGAAAGAAACCAAGTTGCTTAAGATG 1644

```

seq_name: gp_pat:AR098258


```

1401 TGATGCCGCGCTGCTGCGCTTCAGATCTATAGCTGGTGAGCTCCGG 1450
484 1YALVALVALVALLEULYSGLYSGLYSERMETHRGLULYSGLUVAL 500
1451 GAGCTGTTGTTGTTACTTAAGAAAGAAATCTATGACGAAAAAGAGTA 1500
501 MetAspTyrValAlaSerGluValSerAsnAlaLysArgLeuArgGlyGly 517
1501 ATGATATACGTTGCTGCTCAAGTTCATCAATGCAAAACGTTTCGCTGGTGC 1550
517 YVALARGPHEVALASPGIUVALPROLYSGLYLEUTHRGILYSLILEASPG 534
1551 TGTCCGTTTTGGGACGAGTACCTAAAGTCTCACTGTTAAATTCACG 1600
534 1LYLSALAILLEARGSLULLEULYSGLYSERVALALALYSMET 548
1601 GTAAAGCAATTAGAGAAATACGAGAAACACGTTGCTAAGATG 1644

seq_name: gb_pat:E02495
seq_documentation_block:
LOCUS E02495 1644 bp RNA PAT 29-SEP-1997
DEFINITION cDNA encoding luciferase.
ACCESSION E02495
VERSION E02495.1 GI:2170725
KEYWORDS JP 1990171189-A/1.
SOURCE
ORGANISM Luciola lateralis.
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Elateriformia; Cantharoidea; Lampyridae; Luciola.
REFERENCE
1 (bases 1 to 1644)
AUTHORS Kajiyama, N., Tatsumi, H. and Nakano, E.
TITLE LUCIFERASE GENE
JOURNAL Patent: JP 1990171189-A 1 02-JUL-1990;
KIRKMAN CORP
COMMENT
OS Luciola lateralis
PN JP 1990171189-A/1
PD 02-JUL-1990
PE 22-DEC-1988 JP 1988322029
PI KAJIYAMA NAKOKI, TATSUMI HIROKI, NAKANO EICHI PC
C1N15/53//C12N9/02;
CC strandedness: Single;
CC topology: linear;
CC hypothetical: No;
CC anti-sense: No;
FH key Location/Qualifiers
FT mat_peptide 1..1644
FT /product='luciferase'.
FEATURES
source Location/Qualifiers
1..1644
/db_xref='taxon:7052'
BASE COUNT 529 a 262 c 349 g 504 t
ORIGIN
alignment_scores:
Quality: 2814.00 Length: 548
Ratio: 5.144 Gaps: 0
Percent Similarity: 99.818 Percent Identity: 99.635
alignment_block:
US-09-581-241-4 x E02495 ..
Align seg 1/1 to: E02495 from: 1 to: 1644
1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProph 17
1 ATGGAAGAAACATGAGAGAACGATGAAATATTTGTATGTCTGACACCAT 50
17 eTyrProIleGluGluSerAlaGlyAlaGlnLeuArgLysTyrMet 34

```

```

|||||
51 TTACCCATTGGAAGAGGATCTCTGAGACACATTCGCCAGTATATG 100
34 sPaRqTYrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 ATGATATACGAAACTTGAGAGCAATTCCTTTTACTTAACGCACTTACCGGT 150
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGlyG 67
151 GTCGATTATACGATGCGGATCTTAGAAAAATCAATGCTGTCTAGAGAG 200
67 uAlaLeuLysAsnTyrGlyLeuValAlaAspGlyArgGlyIleAlaLeuCys 84
201 GCGCTTAAAGCAATTAATGCTTGGTGTGATGAGAGATTCGCTTATGCA 250
84 eRgLuAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
251 GTGAACACTGGAAGATTCCTTATTCCTGATTAAGCGGTTTATTTATA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
301 GGTGTCGCTGGCTCCAACTAATGAGATTTACACTTACGTGCAATTCGT 350
117 LHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysLysG 134
351 TCACAGTTTAGCATCTCTAAGCCACAAATGATTTAGTTTAAAG 400
134 1YLeuAspLysValIleThrValGlnLysThrValIleThrAlaIleLysThr 150
401 GATTAGATAAGATTATACCTGTACAAAAACGGTACTGCTATTAAAAAC 450
151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAs 167
451 ATTGTTATATTTGACAGACGAAGTGATATATAGAGTTATCATCATCGA 500
167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPhe 184
501 CAACTTTATTAAAAAAACACTCCCAAGGTTTCAAGGATCAAGATTTA 550
184 YSThValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
551 AAACGTGTAAGATTACCGGCAGAAAGACAAAGTTCCTTATATGAACTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLe 217
601 TCGGGTCAACCGGTTTCCAAAGGTGCCAACTTACTCATGAAATGTC 650
217 uValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
651 AGTCACTAGATTTTCTCAAGCTAGAGATCAATTTATGAAACCAAGTTT 700
234 ePrProGlyThrAlaIleLeuThrValAlaProPheHisHisGlyPheGly 250
701 CACACGACGACGCTATTTTAACGTAGTACCATTCACATGAGGTTTGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
751 ATGTTTACTACTTAGGCTATCTAATCTGTGTTTCCGATTTGTCATGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
801 AACGAAATTTGACGAGAGACATTTTAAAAACATCGCAAGATTCAAT 850
284 YSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTTCAAGCGTTATTCCTTACCGACATTCGTTGCAATCTTAAATGAGAGT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
901 GAATTAATCGATTAATATATATTAATCAAAATTAAGTTGAATTCGCTGG 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
|||||

```

COMMENT	OS	KIRKOMAN CORP
PN	Luciola lateralis	
PD	JP 1993244942-A/2	
PF	22-SEP-1993	
PR	22-MAY-1992 JP 1992131057	
PC	27-JUN-1991 JP 91P 157117, 29-NOV-1991 JP 91P 317064	
PC	KAJIYAMA NAOKI, NAKANO EICHI	
PC	C12N9/02, C12N15/53//C12N1/21, (C12N9/02, C12N1:19); CC	
strandedness:	Double;	
CC	topology: linear;	
CC	hypothetical: No;	
CC	anti-sense: No;	
FT	Key	Location/Qualifiers
FT	CDS	1..1644
FEATURES	Location/Qualifiers	/product="Luciola lateralis luciferase"
source	1..1644	
BASE COUNT	529 a 262 c 349 g 504 t	
ORIGIN		
alignment_scores:	quality: 2814.00	length: 548
	ratio: 5.144	gaps: 0
Percent Similarity:	99.818	Percent Identity: 99.635
alignment_block:	US-09-581-241-4 x E05448 ..	
Align seg 1/1	to: E05448	from: 1 to: 1644
1	MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProph	17
1	ATGGAAACATGAGAGACGATCAAAATTTGTGTATGCTCGTAACCAT	50
17	eTyrProIleGluGluGlySerAlaGlyAlaGluLeuArgLysTyrMet	34
51	TTTCCCTAATTGAAGAGGGATCTGCTGGAGCAATTCGCAAGATATGG	100
34	sprArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly	50
101	ATGGAATATGCAAAACCTGGAGCAATTCCTTTACTTAACGCACTTACCGGT	150
51	ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGly	67
151	GTCGATTATACGTACGCCGAATCTTAGAAAAATCATGCTGTAGGAG	200
67	ValLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys	84
201	GCGTTTAAAGAAATTATGGTTGGTTGTGATGAGGAAGAAATTCGTTTGA	250
84	ergLysAsnCysGluLuphePheIleProValLeuAlaGlyLeuPheIle	100
251	GTCGAAACATGTCAGAAATTTCTTATCTCGATTAGCGGTTATTATA	300
101	GlyValGlyAlaAlaProThrAsnGluIleTyrThrLeuArgGluLeu	117
301	GGGTGCGGTGGCTCCCACTAAAGAAATTTTAACTCTACGCAATTCGT	350
117	IHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysG	134
351	TCACAGTTTAGGCATCTCTAAGCCAAATTTGATTAGTTCTTAAAAAG	400
134	LysLeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr	150
401	GATTGATTAAGATTAACTGACAAAAAAGCGTAACTGCTATTAAAAAC	450
151	IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet	167
451	ATTGTTAATTGGACGCAAGGTGGATTATAGAGTTATCAATCATGGA	500

167 pasnheilellysAsnThrProGlnGlyPheLysGlySerPheL 184
 |||||
 501 CAACCTTTATTAATAAAACACCTCCACAGGTTTCAAGATCAAGTTTAA 550
 184 ysthrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
 |||||
 551 AAACGTGAGAACTTAACCGCAAGCAAGTTCCTTATATATGAACCT 600
 201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGlnAsnLe 217
 |||||
 601 TCGGGTTCAACCGGTTTCCCAAAAGTGTCACACTTACTCATGAATAATGC 650
 217 uValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
 |||||
 651 AGTCACTAATTTTCCACAGTAGAGATTCATTTATGGAACCAAGTTT 700
 234 ePProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
 |||||
 701 CACACAGCGACGGCTATTTTAACTGATGATCCATTCATCATGTTTGGT 750
 251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
 |||||
 751 ATGTTTACTACTTATAGCTATCTACTACTGTGTTTTCGATTTGTCATGTT 800
 267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
 |||||
 801 AACGAATTTGACGAGAGACTTTTAAACACACGCAAGATTACAAAT 850
 284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
 |||||
 851 GTTCACAGCGTATTTCTTGTACCGACTTGTGTTCCATTTCTTAATGAACT 900
 301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
 |||||
 901 GAATTAAGTGAATATGATGATTAATCAAAATTGATGAATTCACATCTCG 950
 317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
 |||||
 951 CGGAGCACCTTTATCTMAAAGAAATGGTGAGAGCTGTTGCTAGACGTTTAA 1000
 334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGlnThrThrSerAla 350
 |||||
 1001 ATTACCGGGTTCGTCAGAGCTATGCTTAACAGAAACACCTCTGCA 1050
 351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
 |||||
 1051 ATTATATATCACACCGAGAGCGATGATMAACACAGTCTCTGCAAAAGT 1100
 367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
 |||||
 1101 TGTGCCATTTATTTAAACCAAAAGTTATCGATCTTGATACATAAAACCTT 1150
 384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
 |||||
 1151 TGGCCCGCAACAGACGTGAGAAAGTTGTGTAAAGGCTCTATGCTTATG 1200
 401 LysGlyLysValAspAsnProGluAlaThrArgGluIleIleAspGluG 417
 |||||
 1201 AAAGTTATGATGATTAATCCAGAGCAACAGAAATTCATAGATGAGA 1250
 417 uGlyTyrPheLysHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
 |||||
 1251 AGGTTGGTTGACACAGAGATATTTGGTATTTACAGAGAAACAAACATT 1300
 434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
 |||||
 1301 TCTTATATCGTGAATCGTTTGAAGCTTTTAAATCAATACAAAGGATATCAA 1350
 451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
 |||||
 1351 GTACACCTGCTGATTAATGAAATCTGTCTTTTGCACACATCCAAATATTTT 1400

467 eAspAlaIleValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
 |||||
 1401 TGATGCCGGCGTTCGCTGCGTCCAGATCTTATAGCTGGAGACTCCGG 1450
 484 lValAlaValIleValLeuLysGlyLysSerMetThrGlyLysGluVal 500
 |||||
 1451 GAGCTGTTGTTGTTACTGAAAAAGGAAATCTATGACTGAAAAAGATTA 1500
 501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyG 517
 |||||
 1501 ATGATTACGTTGCTAGTCAAGTTTCAATGCAAAACGTTTGGCTGGTGG 1550
 517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
 |||||
 1551 TGTCCGTTTGTGACAGACGACTTAAGGCTCAGCTGTAATAATTGACG 1600
 534 lLysValAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
 |||||
 1601 GTAAAGCAATTAGAGAAATACTGAAAGAACCGATTGCTAAGATG 1644
 seq_name: gb_in:LLDUCI

seq_documentation_block: 1781 bp mRNA
 LOCUS LLDUCI
 DEFINITION L. lateralis luciferase mRNA.
 ACCESSION X66919.38389
 VERSION X66919.1 GI:9526
 KEYWORDS luciferase
 SOURCE Japanese firefly.
 ORGANISM L. lateralis
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 Elateriformia; Cantharoidae; Lampyridae; Luciola.

REFERENCE
 AUTHORS Tatsumi, H., Kajiyama, N. and Nakano, E.
 TITLE Molecular cloning and expression in *Escherichia coli* of a cDNA
 encoding luciferase of a firefly, *Luciola lateralis*
 JOURNAL Biochim. Biophys. Acta 1131 (2), 161-165 (1992)
 MEDLINE 92305054
 REFERENCE 2 (bases 1 to 1781)
 AUTHORS Tatsumi, H.
 TITLE Direct Submission
 JOURNAL Submitted (21-JUL-1992) H. Tatsumi, Research and Development
 Division, Kikkoman Corporation, 399 Noda, Noda City, Chiba 278,
 JAPAN

FEATURES
 1. 1781 location/Qualifiers
 /organism="Luciola lateralis"
 /db_xref="taxon:7052"
 59..1705
 /codon_start=1
 /product="luciferase"
 /protein_id="CAA47358.1"
 /db_xref="GI:9527"
 /db_xref="SWISS-PROT:O01158"
 /translation="MEMMENDENTVVGPEPEFYPIEGSGAOLRKYMDRYAKGAIATF
 GNALFGDYTYIAEILKESCCLEALIKYGLVNGRLALSCNCEEFIPVLALGTFV
 TNAFTNETYTLRELVHSLGSKPTIVSSKKGLDKTYTQKTAIKTITVILDSKDY
 RYOSMONEIKRNTPOGKSSSEFTVNNRKOVALIMNSGSTGLPKQOLTHENAV
 TRFSHARDPIYGNQVSPGTAITLVPPHNGEMTTGLYLTCGRIVMLTKFDEETFL
 KTDQDCSSVILPTLFAILINRSELDKDLNMLVEIASGGAPLSKEIGEAARREN
 LKGVROGGLRETSALIIIPEDGDKGASGVVPLFKAKVIDLDTKKYIGAPNRGCV
 CVKRPMLMKGVNDPEATRETIIDEGMLHNGDIDGYDEEKHFFIVDLKSLIKYKGV
 VPPAELESVLQHNINFDAGYAGVPPPIAETLPAAVVLKSGSMTEKEVMDVVASGV
 SNARKLRGVFVDEVPKGLKIDGKIRILKRPYAKK"

CDS

BASE COUNT 579 a 276 c 367 g 559 t
 ORIGIN
 alignment_scores: Length: 548
 Quality: 2814.00 Gaps: 0
 Ratio: 5.144
 Percent Similarity: 99.818 Percent Identity: 99.635

alignment_block:
us-09-581-241-4 x LLLUCI ..

Align seg 1/1 to: LLLUCI from: 1 to: 1781

```

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17
  |||||
59 ATGGAACATGAGACGATGCAAAATATGTGTATGCTCTGAACCAT 108
  |||||
17 eTyrProIleGluGluGlySerAlaGlyAlaGluLeuArgLysTyrMet 34
  |||||
109 TTACCTTATGAGAGGATCTGCTGGAGCAATGCGCAAGATATATGG 158
  |||||
34 sPArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
  |||||
159 ATCGATATGCAAAACTTGGAGCAATGCTTTACTAACGCACTTACCGGT 208
  |||||
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysLeuGlyGly 67
  |||||
209 GTCCATTATACGTACGCCGAATCTTAGAAAAATCATGCTGTATAGGAGA 258
  |||||
67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuGly 84
  |||||
259 GCGTTTAAAGAAATATATGTTGTTGTTGATGAGAAATTCGTTATGCA 308
  |||||
84 eArgLysAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
  |||||
309 GTGAAACTGTGAAGAAATCTTTATCTGTATATAGCCGGTTATATTATA 358
  |||||
101 GlyAlaGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
  |||||
359 GGTTGCGGTGTTGCTCCAACTAATGATTTACCTCTACGTGAATGGT 408
  |||||
117 LHisSerLeuGlyLysSerLysProThrIleValPheSerSerLysLys 134
  |||||
409 TCACAGTTTAGCATCTCTAAGCAACAATTGATTTAGTTCAAAAAAG 458
  |||||
134 LysLeuAspLysValIleThrValGluLysThrValThrAlaIleLysThr 150
  |||||
459 GATTAGATTAAGTTATTAAGTGTACAAACAAAGGTTACTGCTATTTAAAC 508
  |||||
151 IleValIleLeuAspSerLysValAspTyrArgLysGlyTyrGlnSerMet 167
  |||||
509 ATGTGATATATGGACAGCAAAAGTGATTTATAGAGTTATCAATCCATGGA 558
  |||||
167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheI 184
  |||||
559 CAACCTTATTAAAAACACCTCCACAGAGTTTCAAGGATCAAGTTTAA 608
  |||||
184 yThrValGluValAsnArgLysGluIleValAlaLeuIleMetAsnSer 200
  |||||
609 AACCTGTAGAGTTAACCCCAAGAACCAAGTTGCTTTTATATGAACCTCT 658
  |||||
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLe 217
  |||||
659 TCGGGTTTCAACCGTTGGCCAAAAGGTGTCAACTTACTCATGAATAATGC 708
  |||||
217 uValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
  |||||
709 AGTCACATAATTTTCTCAGCGTAGAGATCAATTTATGGAACCAAGTTT 758
  |||||
234 eArgProGlyThrAlaIleLeuThrValAlaProPheHisHisGlyPheGly 250
  |||||
759 CACACAGGCGCGGCTATTTAACTAGTACATTCATTCATCATGTTTGGT 808
  |||||
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
  |||||
809 ATGTTTACTACTTTAGGCTATCTAATCTGTGTTTTCGATTTGTCATGTT 858
  |||||
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
  |||||
859 AACGAAATTTGACGAAGACCTTTTAAAAACACTGCAAGATTACAAAT 908
  |||||

```

```

284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
  |||||
909 GTTCAACCGTTATTTCTGTACCGACTTTGTTTCAATCTTTATATCAACT 958
  |||||
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGly 317
  |||||
959 GAATTTACTCATTAATATGATTTATCAAAATTTAGTGAATTCATCTCG 1008
  |||||
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
  |||||
1009 CCGAGACACCTTTATCTAAAGAAATTCGTGAGCTGTGCTAGACGTTTAA 1058
  |||||
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
  |||||
1059 ATTTACCGGGTGTTCGTCAAGGCTATGCTTTAACGAAACCAACTCTGCA 1108
  |||||
351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
  |||||
1109 ATATTATATCACACCGAAGCGCATGATTAACCAAGGTGCTTGGCAAGT 1158
  |||||
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
  |||||
1159 TGTGCATATTATTTAAAGCAAAAGTTATCGATCTGTGATCTAATAAAACTT 1208
  |||||
384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
  |||||
1209 TGGCCCCGAAACAGACGTGGAGAGTTGTGTAAAGGCTCTATAGCTTATG 1258
  |||||
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGluGly 417
  |||||
1259 AAAGGTTATGTAGATTAATCCAGAACCAACAGAAATCATAGATGAAGA 1308
  |||||
417 uGlyThrPheHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
  |||||
1309 AGCTTGTTGTCACACAGAGATTTGGTATTACGATCAAAAAAACACTT 1358
  |||||
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
  |||||
1359 TCTTTATCGTGATGCTTTGCAAGTCTTTATCAATACAAATCAAAAGATATCA 1408
  |||||
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
  |||||
1409 GTACCACTGCTGAATTAAGATCTGTTCTTTGCAACATCCAAATATTTT 1458
  |||||
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
  |||||
1459 TGATCCGCGGCTGTGCTGCGCTCCAGATCTATAGCTGTGAGCTTCGG 1508
  |||||
484 lValAlaValValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
  |||||
1509 GAGCTGTGTTGTACTTGAAGAAAGAAATCTATGACTGAAAAAAGAGTA 1558
  |||||
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyGly 517
  |||||
1559 ATGCTATTAAGTGTGCTAGTCAAGTTTCAAAATGCAAAAGTTTGGTGTGG 1608
  |||||
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
  |||||
1609 TGTCGCTTTTGTGACAGAGTACTTAAGGTCTACACGTGTAATAATTCAGC 1658
  |||||
534 LysAlaIleLeuArgGluIleLeuLysLysProValAlaLysMet 548
  |||||
1659 GTAAAGCAATTAGAGAAATCTGAAAGAAACAGTGTCTAGATG 1702
  |||||
seq_name: gb_pat:AR043317
seq_documentation_block:
LOCUS AR043317 1908 bp DNA
DEFINITION Sequence 8 from patent US 5814465.
ACCESSION AR043317
VERSION AR043317.1 GI:5964325
KEYWORDS

```

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1908)
AUTHORS Matsui, H., Fukuda, S., Kikuchi, M. and Koyama, Y.
TITLE Biotinased firefly luciferase, a gene for biotinased firefly luciferase, a recombinant DNA, a process for producing biotinased luciferase and a bioluminescent analysis method
JOURNAL Patent: US 5814465-A 8 29-SEP-1998;
FEATURES
source location/Qualifiers
1..1908
BASE COUNT 598 a 329 c 426 g 555 t
ORIGIN

alignment_scores:
Quality: 2814.00 Length: 547
Ratio: 5.144 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.817

alignment_block:
US-09-581-241-4 x AR043317 ..

Align seg 1/1 to: AR043317 from: 1 to: 1908

```

1 MetGluAsnMetGluAsnAspGluAsnIleValIleTyrGlyProGluProph 17
1 ATCGAAGAACATGAGAACATGATAATATGTTGATGGCTGACCAACATT 50
17 eTyrProIleGluGluGlySerAlaGlyAlaGluLeuArgLysTyrMetA 34
51 TTACCTATTTGAAGAGGATCTGCTGAGACACATTTGCCAAGTATATG 100
34 sParGTYrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 ATCGATATGCAAAACTTGAGCAATTTGCTTACTAACGCACCTTACCCTG 150
51 ValAspTyrThrTyrAlaGluTyrIleuGluLysSerCysCysLeuGly 67
151 GTGATATTATCGTACGCCGCAATCTTGAAGAAATCATGCTGCTTAGGACA 200
67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
201 GCCTTTAAAGATTATGTTGGTTGTTGTATGAGAACATTTGCTATMGCA 250
84 eArgLysCysGluGluPhePheIleProValIleuAlaGlyLeuPheIle 100
251 GTGAAACTGTGAAGAGTTCTTATTCCTGTATTACCGGCTTATTTATA 300
101 GLyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVa 117
301 GGTTCCGTTGGCTCCAACTAATGAGATTTACACTCTAGCTGATTTGGT 350
117 lHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysLysG 134
351 TCACAGTTTAGGCATCTCTAAGCCAAACAAATTTGATTGTTCAAAAAAG 400
134 lLysLeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
401 GATTAGATAAAGTTATACGTGACAAAAACGGTAACCTGCTATTAAACC 450
151 lIleValIleLeuAspSerLysValAspTyrArgIlyTyrGlnSerMetAs 167
451 ATTGTTATATTGGACAGCAAGTGATTAAGAGTTATATCAATCATGAGA 500
167 pAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheI 184
501 CAACCTTTTAAAAAAACACTCCACAAAGTTTCAAAAGATCAAGTTTAA 550
184 ySThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
551 AAACGTAGAAAGTTTACCGCAAAAGACAAAGTTGCTTTATATGAACTCT 600

```

```

201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLe 217
601 TCGGTTTCAACCGGTTTGGCCAAAAGGTGTCCAACTTACTCATCAATAATT 650
217 uValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
651 GGTACCGCGTTTCTCAGCCTAGAGATCCAAATTTATGAAACCAAGTTT 700
234 eProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
701 CACGAGCAGCGCTATTTTAACTGTACTACCATTCACATCCATGAGTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
751 ATGTTTACTTACTTACGCTATCTACTTGTGGTTTTCGTTATGTCAGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
801 AACGAAATTTGACGAGAGACTTTTAAAAACACTCCAAAGATTACAAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTTCAAGCGTTATCTGTACCGACTTTGTTTGCATTCCTTAATAGAACT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
901 GAATTACTCGATTAATATGATTATCAAAATTTAGTTAAATTTGCATCTGG 950
317 yGlyValAlaProLeuSerLysGluIleGlyGluValValAlaArgArgPhe 334
951 CGGACACCTTTATCTAAAGAAATTTGGAAAGCTGTGCTAGACGCTTTTA 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
1001 ATTACCGCGGTCTTCGTACAGCTATGTTTACAGAAACAACTCTGCA 1050
351 lIleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
1051 ATTATTATCACACCGAAGCGCATGATTAACAGCTGCTTGTGCAAAAGT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysPhe 384
1101 TGTGCAATTATTAAAGCAAAAGTTATGATCTGATCTGATCAAAAAACTT 1150
384 eGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
1151 TGGGCCCGAAGACAGCTGAGAAAGTTTGTAAAGGCTCATGCTATATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGlu 417
1201 AAAGGTATGTAGATATTCAGAACACAAAGAAATCATATGATGAAGA 1250
417 uGlyThrLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
1251 AGGTGTGTTGCACACAGAGATATTTGGGTATACGTGAGAAAAAATTT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
1301 TCTTTATCGTGAATGCTTGAAGTCTTTATCAAAATACAAAGGATATCAA 1350
451 ValProProAlaGluLeuGlnSerValIleLeuGlnHisProAsnIleP 467
1351 GTACACCTCGCTGAATTAAGATCTGTTCTTTTGCACAACTCAAAATATTT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
1401 TGATGCGCGGTTGCTGGGCTTCCAGATCTTATAGCTGTGAGCTTCCG 1450
484 lValAlaValValAlaLeuLysLysGlyLysSerMetThrLysGluVal 500
1451 GACCTGTGTTGTACTTGAAGAAAGAAATCTATGACTGAAAAAGAAAGTA 1500

```


417 uclytrpleuisthrghlyaspilleglytyrtyraspluglulyshtsp 434
 1251 agcttgcttgacacagacagatattggctattacagacacacacacatt 1300
 434 hephellieValAsparglyLeuysSerleuilelytyrlysglytyrigin 450
 1301 tctttatcgatcgatcgatttgaaagcttttaatacctaatacgaagatataca 1350
 451 ValProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIleph 467
 1351 gTACCACCTGCTGAATAGAAATCTCTCTTGGCAACATCCAAATATATTT 1400
 467 eaSpAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
 1401 TGATGCCGCGCTTCTGCTGCGCTTCAGATCTATACCTGTAAGCTTCGG 1450
 484 lValAlaValAlaValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
 1451 GAGCTGTTCTTACTTGAAGAAAGAAATCATGACTGAAAGAAAGATA 1500
 501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
 1501 ATGGATTACGCTGCTAGTCAACATTCCTCAATGCAAAACGTTGCGGTGG 1550
 517 yValArgpheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
 1551 TGCCGCTTTTGGAGAGAGTACCTAAAGCTTCACCTGTTAAATTTGACG 1600
 534 lLysAlaIleArgGluIleLeuLysLysProValAlaLys 547
 1601 GTAAAGCAATGTAGAAATACCTGAAGAAACGAGTTGCTAAG 1641

seq_name: gb_pat:EI2279

seq_documentation_block:

LOCUS EI2279 1908 bp DNA PAT 24-JUN-1998
 DEFINITION cDNA encoding biotinylated Luciola luciferase.
 ACCESSION EI2279
 VERSION EI2279.1 GI:3251113
 KEYWORDS JP 1996308578-A/4.
 SOURCE
 ORGANISM

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 Elateriformia; Cantharidea; Lampyridae; Luciola.

1 (bases 1 to 1908)
 Tatsumi,H., Fukuda,M., Kikuchi,M. and Koyama,T.

BIOTIN FIRE FLY LUCIFERASE, BIOTIN FIRE FLY LUCIFERASE GENE, NEW
 RECOMBINANT DNA, PRODUCTION OF BIOTIN FIRE FLY LUCIFERASE AND
 BIOLUMINESCENCE ANALYSIS

Patient: JP 1996308578-A 4 26-NOV-1996;
 KIKKOMAN CORP

COMMENT
 OS Luciola lateralis
 PN JP 1996308578-A/4

PD 26-NOV-1996
 PF 24-APR-1995 JP 1995098857

PR 27-JUL-1994 JP 94P 193798, 14-MAR-1995 JP 95P 54625 PI
 TATSUMI HITOKI, FUKUDA MASARU, KIKUCHI MAMORU, KOYAMA TAJUJI PC

C12N15/09,C12N9/02,C12Q1/26,(C12N9/02,C12R1:185); CC
 strandedness: Double;

CC topology: Unknown;
 CC hypothetical: No;

CC anti-sense: No;

FH key location/Qualifiers

FT source 1. 1908 /organism='Luciola lateralis' FT

FT /clone='pHLI248' 1. 1908
 FT mat-peptide /product='biotinylated Luciola luciferase'.

FEATURES
 source 1. 1908
 /organism='Luciola lateralis'

BASE COUNT /db_xref='taxon:7052'
 598 a 329 c 426 g 555 t
 ORIGIN

alignment_scores:
 Quality: 2814.00 Length: 547
 Ratio: 5.144 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.817

alignment_block:
 US-09-581-241-4 x EI2279 ..

Align seg 1/1 to: EI2279 from: 1 to: 1908

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProH 17
 1 ATGGAAAAACATGGAGAACGATGAAAAATATGTGTATGCTCTGAACCAT 50
 17 eTyrProIleGluGluLysSerAlaGlyAlaGluLeuArgLysTyrMet 34
 51 TTACCCATTGTAAGAGGATCTGCTGGACACAATTCGCAAGTATGG 100
 34 spArgTyrAlaLysLeuGlyAlaIleAlaIleAlaIleAlaIleAla 50
 101 ATCATATATGCAAAACCTTGACCAATTCCTTTACTTAACGCACTTACCG 150
 51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysLeuGly 67
 151 GTCGATTATACGACCGCCGAATCTTGAAGAAATCATAGCTGTACAGAGA 200
 67 uAlaLeuLysAsnTyrGlyLeuValAlaSpGlyArgIleAlaLeuCys 84
 201 GGCTTTAAAGAAATATGTTGTTGTTGTTGATGAGAAATTCGTTATGCA 250
 84 eArgLysCysGluGluPhePheIleProValIleuAlaGlyLeuPheIle 100
 251 GTGAACCTGTGAAGAGTCTTATCTCTGATTAACCGGTTATTTATA 300
 101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeu 117
 301 GGTGCGGTGGCTGCCAATATGATGATTAACCTGACTGAATGGT 350
 117 HisSerLeuGlyIleSerLysProThrIleValPheSerSerLysG 134
 351 TCACAGATTAGGACATCTTAAGCCCAATGTATTGTTAGTTTAAAG 400
 134 lLysLeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
 401 GATTGATTAAGTTATTAACGTACAAAACGATACGTAAGCTGCTTAAACC 450
 151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
 451 ATTGTTATATTGGACAGCAAACTGATTAAGAGTTATCAATCATGGA 500
 167 pAsnPheIleLysLysAsnThrProGlnGlyPheLysSerSerPhe 184
 501 CAACCTTATTAATAAAAAACACCCCAAGGTTTAAAGGATCAAACTTTA 550
 184 yThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
 551 AAACGTATGAAGTTAACCGCAAAAGAAAGATGCTCTTATATAGAACCT 600
 201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsn 217
 601 TCGGGTTCAACCGGTTTGCACAAAGGTGCAACTTACTCATCAAGAAATTT 650
 217 uValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnVal 234
 651 GGTACAGCGCTTTTCTCACGCTAGAGATCAATTTATGGAACCAAGTTT 700
 234 eProGlyThrAlaIleLeuThrValAlaProPheHisGlyPheGly 250

```

701 CACGAGCAGCGCTATTTAACTGATACCATTCATCATGTTTGGT 750
251 MetPheThrLeuGlyTyrLeuThrcysGlyPheArgIleValMetIle 267
751 ATGTTTACTTACTTAGGCTATCTACTTGTGTGTTTGTGATGTGCAAGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
801 AACCAAAATTCAGCAGAGACTTTTAAAAAACAACCTGCAAGATTACAAAT 850
284 ysserServalIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTTAAAGCGTTATTCTGTACCGACTTGTGTGCAATTTCTTAATAGAAAT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
901 GAATTACTCGATAATATGATTTATCAAAATTTAGTTGCAATTTGCACTGG 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaIleArgArgPhe 334
951 CGGAGCACCTTTATCTAAAGAAATGGTGAAGCTGTGCTAGACGTTTAA 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrPheSerAla 350
1001 ATTTACCGGGTGTTCGTCAGAGCTATGCTTTAAACAGAAACAACCTCTGCA 1050
351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
1051 ATTATTTATCACCGCGAGAGCGATGATTAACACAGTGCTTCTGCAAAAGT 1100
367 ValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
1101 TGTGCCATTTTAAAGCAAAAGTTATCGATCTGTGATCTAAAAAAAAGCTT 1150
384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
1151 TGGGCCCCGAAACAGACGTCGGAAGTTGTGTAAAGGTCCTATGCTTAAG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGlu 417
1201 AAAGGTATGTATGATATATCCAGAACACACAGAAATCATGATGAAAGA 1250
417 uGlyTyrPheHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
1251 AGGTGTGTTGCACACAGAGATATTTGGTATTAACGATGAAGAAAAACATT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
1301 TCTTATTCGTGATCGTTGAAGCTTTATATCAAAATACAAAGATATATCA 1350
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePhe 467
1351 GTACCACTGCTGCAATTTAGAAATCTGTTCTTTGCAACATCCAAATATTTT 1400
467 eaSpAlaGlyValAlaGlyValProAspProIleAlaGlyLysLeuProG 484
1401 TGCATGCCGCGTGTGCGTCCAGATCTTACTGCTGAGCTTCCGG 1450
484 yAlaValValIleLeuLysLysGlyLysSerMetThrGluLysGluVal 500
1451 GAGCTGTGTTGTTACTGAAAAAGAAAAATCTATGACTGAAAAAGAAAGTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
1501 ATGATATACGTTGCTAGTCAATGTTCAAAATGCAAAACGTTTGGCGGGGG 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysLysLeaspG 534
1551 TGTCCGTTTGTGGAGAGTACTAAAGGCTCTCACTGTAATAATTTGACG 1600
534 yLysAlaIleArgGluIleLeuLysLysProValAlaLys 547
1601 GTAAAGCAATTAGAGAAATACTGAGAAACACAGTGTAAAG 1641

```

```

seq_name: gb_pat:E16288
seq_documentation_block:
LOCUS      E16288          1920 bp      DNA          PAT          28-JUL-1999
DEFINITION DNA encoding fusion protein which comprises luciferase and
            lysostaphin.
ACCESSION  E16288
VERSION    E16288.1  GI:5710971
KEYWORDS   JP 1998150991-A/1.
SOURCE     unidentified.
ORGANISM   unidentified
            unclassified.
REFERENCE  1 (bases 1 to 1920)
AUTHORS   Tatsumi,H., Fukuda,M. and Nagahara,A.
TITLE      LUCIFERASE-LYOSTAPHA IN FUSED PROTEIN, ITS PRODUCTION AND
            BIOLUMINESCENT ANALYSIS USING THE SAME
JOURNAL    Patent: JP 1998150991-A 1 09-JUN-1998;
            KIKKAWAN CORP
COMMENT    OS   None
            OC   Artificial sequences.
            PN   JP 1998150991-A/1
            PD   09-JUN-1998
            PF   25-NOV-1996  JP 1996328042
            PI   TATSUMI HIROKI, FUKUDA MASARU, NAGAHARA AYUMI PC
            CI  1/2N15/09,C07H21/04,C07K19/00,C12N9/02,C12N9/52,C12P21/02,PC
            CI  201/66
            PC   (C12N9/02,C12R1:19),(C12N9/52,C12R1:19);
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            FH   Key
            FT   Location/Qualifiers
FEATURES             Location/Qualifiers
     source           1..1920
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT          627 a 310 c 407 g 576 t
ORIGIN
alignment_scores:
    Quality: 2814.00      Length: 547
    Ratio: 5.144          Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 99.817
alignment_block:
US-09-581-241-4 x E16288 ..
Align seg 1/1 to: E16288 from: 1 to: 1920
1 MetLusMetGluAsnAspGluAsnIleValTyrGlyProGluProH 17
|||||
1 ATGAAAAACATGAGAGAGATGAAAAATATGTGTATGCTCAACCAT 50
17 eTyrProIleGluGluSerAlaGluIleAlaPheThrAsnAlaLeuThrGly 34
|||||
51 TTACCCATTGGAAGAGGATCTGCTGGAGCACAATTCGCGCAATATATG 100
51 TACCATATGCAAAACTGGAGCAATTCGTTTCTTAACGACTTACCGGT 150
34 spArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
|||||
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGly 67
|||||
151 GTCGATTATACGTAACGCCGCAATCTTAGAAAAATCAATGCTGTCTAGAGA 200
67 uAlaLeuLysAsnTyrGlyLeuValAlaAspGlyArgIleAlaLeuCys 84
|||||
201 GCCTTAAAGAAATATGCTTGTGTTGATGGAAGAAATTCGTTATGCA 250

```

```

84  erGlusnCySGluGluNhepHeileProvalleuAlaGlyLeuPheile 100
|||||
251  GTGAAGACGTGAAGAGTTCTTATCTGATTAAGCCGGTTATTTATA 300
|||||
101  GtYValGtYValAlaProthrasGluileThyrThreleuArGluLeuVa 117
|||||
301  GGTGCGGTGGCTCCACTAATGAGATTACACTGACGTGGAATGGT 350
|||||
117  LHiserleuGlyIleSerLysProthrIleValPheSerSerLysLysG 134
|||||
351  TCACAGTTAGGCATCTCTAAGCCACAACTGTATTAGTTCTAAAAAG 400
|||||
134  lYleuAspLysValIleThrValGlnLysThrValThAlaIleLysThr 150
|||||
401  GATTAGATTAAAGTTATTAAGTACAAAAACGGTACCTGATTAAAAAC 450
|||||
151  lIleValIleleuAspSerLysValAspLysPyrArgLysIleSerMetAs 167
|||||
451  ATTGTTATATTGGACACGAAAGTATTAAGAGTTATCAATCCATGGA 500
|||||
167  pasnPhelIeLysLysAsnThrProGlnGlyPheLysSerSerPheL 184
|||||
501  CAACCTTTATTTAAAAAACACATCCACAAGGTTTCAAGGATCAAGTTTA 550
|||||
184  ysthrValGluValasnArgLysGluGlnValAlaLeuileMetLysnser 200
|||||
551  AAACGTGTAGAGTTAACCGCAAGAACAAAGTGTCTTTATTAATGAACTCT 600
|||||
201  SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluasnle 217
|||||
601  TCGGGTTCAACCGGTTTGGCAAAAGGTGCAAACTTACTCATGAAATTT 650
|||||
217  uValThrArgPheSerHisAlaArgAspProIleThyGlyAsnGlnValS 234
|||||
651  GGTCAAGCGCTTTTTCACGCTAGAGATCAATTTATGAAACCAAGTTT 700
|||||
234  expProGlyThrAlaIleleuThrValAlaProPheHisHisGlyPheGly 250
|||||
701  CACGAGGACGCGTATTTTAACTGATGATCCATTCATCATGTTTGGT 750
|||||
251  MetPheThrThrLeuGlyTyThrLeuThrcysGlyPheArgIleValMetle 267
|||||
751  ATGTTTACTACTTATGAGTATCTAATCTGGTTTGGTATTTGTCATGTT 800
|||||
267  uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspLysLysC 284
|||||
801  AACGAAATTTGACGAAGAGACTTTTAAAAACATGCAAGATTACAAAT 850
|||||
284  ySerSerValIleleuValProThrLeuPheAlaIleleuAsnArgSer 300
|||||
851  GTTCAGAGCGTTATCTGTATACGACGCTTTGTTTCAATTCCTTATAGAG 900
|||||
301  GluLeuLeuAspLysLysAspLeuSerAsnLeuValGluIleAsnArgI 317
|||||
901  GAATTACTCGATAAATATGATTTATCAAAATTTGTTGAATTTGCACTCTG 950
|||||
317  yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
|||||
951  CGGAGCACCTTTTATCTTAAGAAATTTGTTGAAGCTGTTGCTAGACGTTTA 1000
|||||
334  snLeuProGlyValArgGlnGlyTyThrLeuThrcysGlyLeuThrSerAla 350
|||||
1001  ATTTACCGGGTGTTCGTCAGAGCTATGTTTAAACAGAAACAACTCTGCA 1050
|||||
351  lIleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
|||||
1051  ATTATTTATCACACCGAAGCGATGATTAACCAAGGTCTTCTGCAAGT 1100
|||||
367  lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
|||||
1101  TGTGCAATTTATTAAGCAAAAGTTATCGATCTTGATACATAAAAAAACTT 1150
|||||

```

```

384  euGIYProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
|||||
1151  TGGGCCCGAACACAGACGTGGAGAACTTTGTCTAAAGGCTCTATAGCTTAG 1200
|||||
401  LysGlyTyThrValAspAsnProGluAlaThrArgGluIleIleAspGluG 417
|||||
1201  AAAGGTATGATGATTAATCCAGAAAGCAAGAAAGAAATCATGATGAGAG 1250
|||||
417  uGlyTyThrLeuHisThrCysLysPheIleGlyTyThrAspGluGlnLysHisP 434
|||||
1251  AGGTGGTTCACACAGAGATATTTGGTATTTACGATGAAGAAAAACATT 1300
|||||
434  hePheIleValAspArgLeuLysSerLeuIleLysTyThrGlyTyGln 450
|||||
1301  TCTTTATCGTGATCGCTTGAAGCTTTATCAAAATCAAGATATCA 1350
|||||
451  ValProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePhe 467
|||||
1351  GTACCACTGCTGAATTAGAAATCTGTTCTTTGCAACATCCAAATATTTT 1400
|||||
467  easpAlaGlyValAlaGlyValProAspProIleAlaGlyIleuLeuProG 484
|||||
1401  TGATGCCGGCGCTGCTGCGCTTCCAGATCTATAGCTGGTGAGCTTCGG 1450
|||||
484  lYAlaValAlaValAlaLeuLysLysGlyLysSerMetThrGluLysGluVal 500
|||||
1451  GAGCTGTTGTTGTTACTTGAAGAAAAAGAAATCTATGATCAAGAAAAAGTA 1500
|||||
501  MetAspTyThrValAlaSerGlnValSerAsnAlaLysArgLeuArgLysG 517
|||||
1501  ATGATTAAGTTAGTTGCTAGTCAAGTTCAAAATCCAAAGAGTTGCGTGG 1550
|||||
517  yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
|||||
1551  TGTCCGTTTGTGGACGAGATCACTAAAGGTCTCAGCTGGTAAATGACG 1600
|||||
534  lYlYsAlaIleArgGluIleLeuLysLysProValAlaLys 547
|||||
1601  GTAAAGCAATTAGAGAAATTAAGTAAAGAAACCAAGTTGCTTAG 1641
|||||
seq_name: gb_pat:AR043316
seq_documentation_block:
LOCUS AR043316 1704 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 5 from patent US 5814465.
ACCESSION AR043316
VERSION AR043316.1 GI:5964324
KEYWORDS
SOURCE
ORGANISM Unknown.
UNCLASSIFIED.
REFERENCE 1 (bases 1 to 1704)
AUTHORS Tatsumi,H., Fukuda,S., Kikuchi,M. and Koyama,Y.
TITLE Biotinated firefly luciferase, a gene for biotinated firefly
luciferase, a recombinant DNA, a process for producing biotinated
luciferase and a bioluminescent analysis method
JOURNAL Patent: US 5814465-A 5 29-SEP-1998;
FEATURES
source Location/Qualifiers
BASE COUNT 541 a 275 c 361 g 527 t
ORIGIN
alignment_scores:
Quality: 2801.00 Length: 546
Ratio: 5.130 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.451
alignment_block:
US-09-581-241-4 x AR043316 ..
Align seg 1/1 to: AR043316 from: 1 to: 1704

```

3 AsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPheTyrPr 19
 : : : : :
 67 AGTCTCGAAGACATGAAATATGTGTATGTCTCTGACCATTTTACC 116
 19 oileGluGlySerAlaGlyAlaGlnLeuArgLysTyrMetAspArgT 36
 : : : : :
 117 TATGAAGAGGATCTGCTGAGACCAATTCGCAAGATATGATCATGAT 166
 36 yralAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGlyValAsp 52
 : : : : :
 167 ATGCAAAACCTTGAGCAATTCCTTTTACTACACGACTTACCGGTGTGAT 216
 53 TyrThrTyrAlaGlyTyrLeuGlyLysSerCysCysLeuGlyAlaLe 69
 : : : : :
 217 TATAAGTACCCGCAATCTAGAAATAATCATGCTGTAGAGAGGCTT 266
 69 uLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCysSerGlu 86
 : : : : :
 267 AAAGAAATATGCTTGTGTTGTGATGAGAAAGATTCGTTATGACGATAA 316
 86 sncCysGluGluPhePheIleProValLeuAlaGlyLeuPheIleGlyVal 102
 : : : : :
 317 ACTGTGAAGAGTCTTATTCCTGTATATACCGGTTATTTATATAGTGTG 366
 103 GlyValAlaProThrAsnGluIleTyrThrLeuArgLysLeuValHisSe 119
 : : : : :
 367 GGTGTGCTCCAACTAATGAGATTTACACTCTACGTAATGTTGTTACAG 416
 119 rLeuGlyIleSerLysProThrIleValPheSerSerLysLysGlyLeu 136
 : : : : :
 417 TTTAGGATCTCTAAGCCACAAATTTGATTTAGTCTAAATAAAGATTAG 466
 136 spLysValIleThrValGlnLysThrValThrAlaIleLysThrIleVal 152
 : : : : :
 467 ATTAAGTTATTAAGTACAAATAAAGGTTACTGTTATTAATAACCATGTT 516
 153 IleLeuAspSerLysValAspTyrArgLysTyrGlnSerMetAspAsnPh 169
 : : : : :
 517 ATATTGAGACGCAAGTGGATTATAGAGTTATCAATCCATGGACACAT 566
 169 eileLysLysAsnThrProGlnGlyPheLysGlySerSerPheLysThrV 186
 : : : : :
 567 TATTAATAAATAACACTCCACAAAGTTTCAAGATCAAGTTTAAACCTG 616
 186 aLgIleValAsnArgLysGluGlnValAlaLeuIleMetAsnSerSerGly 202
 : : : : :
 617 TAGAAGTTAAACGCAAGAAAGATGCTCTTATTAATGAACCTTTGGGCT 666
 203 SerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLeuValTh 219
 : : : : :
 667 TCAACCGGTTTGGCAAAAGGTGTGCAACTTACATCAATGAATTTGGTCT 716
 219 rArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValSerProG 236
 : : : : :
 717 GCGTTTCTTCACGCTAGAGATCCAAATTTATGAAACCAAGTTTCCACAG 766
 236 lYThrAlaIleLeuThrValValProPheHisHisGlyPheGlyMetPhe 252
 : : : : :
 767 GCACGGCTATTTTAACTAGTACCATTCATCATGCTTTGGTATGTTT 816
 817 ACTACTTAAAGGCTATCTAAGTGTGTTTCTGATTTGTCATGTAAAGAA 866
 269 sPheAspGluGluThrPheLeuLysThrLeuLysAspTyrLysCysSerS 286
 : : : : :
 867 ATTTCACGAGAGACTTTTAAACAACTGCAAGATTAACAATGTTCAA 916
 286 eValIleLeuValProThrLeuPheAlaIleLeuAsnArgSerGluLeu 302
 : : : : :
 917 GCGTATATCTGTACGACTTGTGTTGCAATCTTATATGAAGTAATTA 966

303 LeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGlyVal 319
 : : : : :
 967 CTCGTAATATATGATTATTAATTAATTTAGTGAATTTGCATCTGGCGGAGC 1016
 319 aProLeuSerLysGluIleGlyAlaValAlaArgArgPheAsnLeuP 336
 : : : : :
 1017 ACCTTTATCTAAAGAAATTTGTTGAAGCTGTGTGACAGCTTTTAATTTAC 1066
 336 roGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAlaIleLe 352
 : : : : :
 1067 CGGGTGTTCGCAAGGCTATGTTTAAACAGAAACACCTCTGCAATTAAT 1116
 353 IleThrProGluGlyAspAspLysProGlyValAserGlyLysValAlPr 369
 : : : : :
 1117 ATCAACCCGAGAGCGATGATTAACCAAGTCTTGTGCAAAAGTTGTGCC 1166
 369 oLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGlyP 386
 : : : : :
 1167 ATTAATTAAGCAAAAGTTATGATCTTGATTAATAAATAAATTTGGGCC 1216
 386 roAsnArgArgGlyGlyValCysValLysGlyProMetLeuMetLysGly 402
 : : : : :
 1217 CGAACAGACGTCGAGAGATTTGTGTTAAAGGCTCTATGCTTATGAAAGCT 1266
 403 TyrValAspAsnProGluAlaThrArgGluIleLeuAspGluGlyTyr 419
 : : : : :
 1267 TATGTAGATATATCCAGAACAGACAGAAATCATATGATGAAAGGTTG 1316
 419 pLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisPhePhe 436
 : : : : :
 1317 GTTGACACAGAGAGATTTGGGATTTACGATGAAGAAATAAATTTCTTTA 1366
 436 leValAspArgLysSerLeuIleLysTyrLysGlyTyrGlnValPro 452
 : : : : :
 1367 TCGGATCGCTTTGAAGTCTTTATCAATCAATCAAGATATCAAGTACCA 1416
 453 ProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePheAspAl 469
 : : : : :
 1417 CCGTGTGAATTAAGATCTGTTCTTTGCAACATCCAAATATTTTGTATGC 1466
 469 aGlyValAlaGlyValProAspProIleAlaGlyGluLeuProGlyValAl 486
 : : : : :
 1467 CGGCGTTCGCTGGGCTTCAGATCTTATAGCTGTGAGCTTCCGGGAGCTG 1516
 486 aValValLeuLysLysGlyLysSerMetThrGluLysGluValMetAsp 502
 : : : : :
 1517 TTGTGTACTTGAATAAAGAAATCTATGACTGATAAAGAAAGATTAATGANT 1566
 503 TyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyValAlAr 519
 : : : : :
 1567 TACGTGTAGTCAAGTTCCTCAATGCAAAACGTTTCCGTGTGTGTCCG 1616
 519 gPheValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLysA 536
 : : : : :
 1617 TTTTGGAGAGAAAGTACCTTAAGGTTCTACTGTGTAATTTGACGAGTAAG 1666
 536 lAlaArgGluIleLeuLysLysProValAlaLysMet 548
 : : : : :
 1667 CAATTAGAGAAATACTGAAGAAACAGTTGCTAAGATG 1704

seq_name: gb_pat: AR062708

seq_documentation_block: 1704 bp DNA

LOCUS AR062708 Sequence 5 from patent US 5843746. PAT 29-SEP-1999

DEFINITION AR062708

ACCESSION AR062708

VERSION AR062708.1 GI:5990399

KEYWORDS

SOURCE

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1704)

AUTHORS Tatsumi,H., Fukuda,S., Kikuchi,M. and Koyama,Y.

TITLE

Biotinlated firefly luciferase, a gene for biotinated firefly luciferase, a recombinant DNA, a process for producing biotinated luciferase and a bioluminescent analysis method

JOURNAL

Patent: US 5843746-A 5 01-DEC-1998;

FEATURES

Location/Qualifiers
1..1704

/organism="unknown"

BASE COUNT 541 a 275 c 361 g 527 t
ORIGIN

alignment_scores:

Quality: 2801.00 Length: 546
Ratio: 5.130 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.451

alignment_block:

US-09-581-241-4 x AR062708 ..

Align seg 1/1 to: AR062708 from: 1 to: 1704

```

3 AsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPheTyrPr 19
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
67 AGCTCGAGACGATGATAATATTGATGCTCCGAGACCATTTTACCC 116
19 otleGluGluGlySerAlaGlyAlaGluIleuArgLysTyrMetAspArg 36
   |||TATTTGAAGAGGATCTGCGAGACACAAATTCGCAAGTATATGATGCAT 166
117 TATTTGAAGAGGATCTGCGAGACACAAATTCGCAAGTATATGATGCAT 166
36 TyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGlyValAsp 52
   |||ATGCAAAACTGGAGCAATTCCTTACTACGCACTTACCGGTTGCGAT 216
167 ATGCAAAACTGGAGCAATTCCTTACTACGCACTTACCGGTTGCGAT 216
53 TyrThrTyrAlaGluTyrLeuGluLysSerCysLeuGlyGluAlaIle 69
   |||TATACGTACCGCGAATCTAGAAAAATCATGCTGCTAGAGAGGCTTT 266
217 TATACGTACCGCGAATCTAGAAAAATCATGCTGCTAGAGAGGCTTT 266
69 uLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCysSerGlu 86
   |||TATGATGATGTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 316
267 TATGATGATGTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 316
86 snCysGluGluPhePheIleProValIleuAlaGlyLeuPheIleGlyVal 102
   |||TATGATGATGTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 366
317 TATGATGATGTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 366
103 GlyValAlaProThrAsnGluIleTyrThrLeuArgLysLeuValHis 119
   |||GGTGTGCTCCAACTAATGAGATTTACACTCTACGTGAATGGTTCACAG 416
367 GGTGTGCTCCAACTAATGAGATTTACACTCTACGTGAATGGTTCACAG 416
119 rLeuGlyIleSerLysProThrIleValIlePheSerSerLysGlyLeuAla 136
   |||TTTGGCATCTCTAAGCCACAAATTTGATTTAGTTCTAAAAAAGGATTA 466
417 TTTGGCATCTCTAAGCCACAAATTTGATTTAGTTCTAAAAAAGGATTA 466
136 sPlyValIleThrValGlnLysThrValIleAlaIleLysThrIleVal 152
   |||ATAAGTTAATACGTACAAAAAAGCGTAACTCTATTAATAAACCATTTG 516
467 ATAAGTTAATACGTACAAAAAAGCGTAACTCTATTAATAAACCATTTG 516
153 IleuAspSerLysValAspTyrArgGlyTyrGlnSerMetAspAsnPh 169
   |||ATATTGGACGCAAGTGAATTATAGAGTTATCAATCCATGGACAACCT 566
517 ATATTGGACGCAAGTGAATTATAGAGTTATCAATCCATGGACAACCT 566
169 eileLysLysAsnThrProGlnGlyPheLysGlySerSerPheLysThrV 186
   |||TATTTAAAAAAACACTCCACAGTTTCAAAAGATCAAGTTTAAAACTG 616
567 TATTTAAAAAAACACTCCACAGTTTCAAAAGATCAAGTTTAAAACTG 616
186 alGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSerSerGly 202
   |||TAGAAGTTAACCGCAAAAGAACAGTTGCTCTTATTAATGAACCTCTT 666
617 TAGAAGTTAACCGCAAAAGAACAGTTGCTCTTATTAATGAACCTCTT 666
203 SerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLeuValTh 219
   |||TCAACGGGTTGGCCAAAAGGTGTGCAACTTACTCATGAAAAATTTGGT 716
667 TCAACGGGTTGGCCAAAAGGTGTGCAACTTACTCATGAAAAATTTGGT

```

```

219 rArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValSerProG 236
   |||GCGTTTTCACCTTGAGATCCAAATTTATGAAACCAAGTTTCACAG 766
717 GCGTTTTCACCTTGAGATCCAAATTTATGAAACCAAGTTTCACAG 766
236 lYThrAlaIleLeuThrValValProPheHisHisGlyPheGlyMetPhe 252
   |||GCACGGCATTTTAACTGATGATACATTCATCATGATTTGGTATGTTT 816
767 GCACGGCATTTTAACTGATGATACATTCATCATGATTTGGTATGTTT 816
253 ThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLeuThrLy 269
   |||ACTACTTAGGCTATCTACTGTTGGTGTGTGTGTGTGTGTGTGTGTGT 866
817 ACTACTTAGGCTATCTACTGTTGGTGTGTGTGTGTGTGTGTGTGTGTGT 866
269 sPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysCysSer 286
   |||ATTGACCAAGACACTTTTAAAAACACGCAAGATTAACAATGTTCAA 916
867 ATTGACCAAGACACTTTTAAAAACACGCAAGATTAACAATGTTCAA 916
917 GCGTTATCTCTGTACCGCACTTGTTCAAATCTTAATAGAAATGAAATTA 966
286 erValIleLeuValProThrLeuPheAlaIleLeuAsnArgSerGluLeu 302
   |||GCGTTATCTCTGTACCGCACTTGTTCAAATCTTAATAGAAATGAAATTA 966
917 GCGTTATCTCTGTACCGCACTTGTTCAAATCTTAATAGAAATGAAATTA 966
303 LeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGlyVal 319
   |||CTCGATTAATATGATTTATCAAAATTTAGTTGAATTCGATCTGGCGAGC 1016
967 CTGATTAATATGATTTATCAAAATTTAGTTGAATTCGATCTGGCGAGC 1016
319 aProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheAsnLeuP 336
   |||ACCTTATCTTAAGAAATTTGGTGAAGCTGTTGCTAGACGTTTAAATTTAC 1066
1017 ACCTTATCTTAAGAAATTTGGTGAAGCTGTTGCTAGACGTTTAAATTTAC 1066
336 roGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAlaIleIle 352
   |||CGGGTGTGCGCAAGCTATGCTTTAAACAAACAACTCTGCAATTAAT 1116
1067 CGGGTGTGCGCAAGCTATGCTTTAAACAAACAACTCTGCAATTAAT 1116
353 lIeThrProGluGlyAspAspLysProGlyAlaSerGlyLysValValPr 369
   |||ATCACACCGCAAGCGGATGATTAACACAGCTCTCTGCAAAAGTTGGCC 1166
1117 ATCACACCGCAAGCGGATGATTAACACAGCTCTCTGCAAAAGTTGGCC 1166
369 oLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGlyP 386
   |||ATTATTTAAAGCAAAAGTTATCGATCTTGATCACTAAAAAACTTTGGGCC 1216
1167 ATTATTTAAAGCAAAAGTTATCGATCTTGATCACTAAAAAACTTTGGGCC 1216
386 roAsnArgArgGlyGluValCysValLysGlyProMetLeuMetLysGly 402
   |||TATGATGATGTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1266
1217 TATGATGATGTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1266
403 TyrValAspAsnProGluAlaThrArgLysIleIleAspGluGlyTyr 419
   |||TATGATGATGTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1316
1267 TATGATGATGTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1316
419 pleuHisThrGlyAspIleGlyTyrTyrAspGluLysHisPhePheI 436
   |||TATGATGATGTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1366
1317 TATGATGATGTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1366
436 lValAlaPargLeuLysSerLeuIleLysTyrLysGlyTyrGlnValPro 452
   |||TATGATGATGTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1416
1367 TATGATGATGTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1416
453 ProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePheAspAl 469
   |||CTGCTGAATTAAGATCTGCTTTTGGCAACATCCAAAATTTTGGATGC 1466
1417 CTGCTGAATTAAGATCTGCTTTTGGCAACATCCAAAATTTTGGATGC 1466
469 aglyValAlaGlyValProAspProIleAlaGlyGluLeuProGlyAlaV 486
   |||GCGGTTGGTGGGCTTCCAGATCTTATACCTGCTGAGCTTCGGGAGCTG 1516
1467 GCGGTTGGTGGGCTTCCAGATCTTATACCTGCTGAGCTTCGGGAGCTG 1516
486 alValValLeuLysLysGlyLysSerMetThrGluLysGluValMetAsp 502
   |||TTTGTGTACTTGAAAAAGGAAATCTATGACTGAAATAAGATATGAT 1566
1517 TTTGTGTACTTGAAAAAGGAAATCTATGACTGAAATAAGATATGAT 1566
503 TyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyGlyValAr 519
   |||TACGTTGCTAGTCMAAGTTCAAAATGCAAAACGTTTCGGGTGCTCG 1616
1567 TACGTTGCTAGTCMAAGTTCAAAATGCAAAACGTTTCGGGTGCTCG 1616
519 gpHeValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLysA 536

```

```

|||||
1617 TTTTGTGGACGAGTAACCTPAAGTCTCAGTGAATAATTGACGGTAAG 1666
536 laileargluileuylslyspovalalalysmet 548
1667 CAATTAGAGAAATFACTGAGAAACAGTTGCTAAGATG 1704
seq_name: gb_pat:E12278

seq_documentation_block:
LOCUS E12278 1704 bp DNA PAT 24-JUN-1998
DEFINITION cDNA encoding biotinylated Luciola luciferase.
ACCESSION E12278
VERSION E12278.1 GI:3251112
KEYWORDS JP 1996308578-A/3.
SOURCE Luciola lateralis.
ORGANISM Luciola lateralis.
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
Elateriforma; Cantharoidae; Lampyridae; Luciola.
REFERENCE 1 (bases 1 to 1704)
AUTHORS Tatsumi,H., Fukuda,M., Kikuchi,M. and Koyama,T.
TITLE BIOTIN FIRE FLY LUCIFERASE, BIOTIN FIRE FLY LUCIFERASE GENE, NEW
JOURNAL RECOMBINANT DNA, PRODUCTION OF BIOTIN FIRE FLY LUCIFERASE AND
KIRKMAN CORP. BIOLUMINESCENCE ANALYSIS
PATENT: JP 1996308578-A 3 26-NOV-1996;
COMMENT KIRKMAN CORP
OS Luciola lateralis
PN JP 1996308578-A/3
PD 26-NOV-1996
PF 24-APR-1995 JP 1995098857
PR 27-JUL-1994 JP 94P 193798, 14-MAR-1995 JP 95P 54625 PI
TATSUMI HIROKI, FUKUDA MASARU, KIKUCHI MAMORU, KOYAMA TADUO PC
C12N15/09,C12N9/02,C12O1/26,C12N9/02,C12R1/185; CC
Strandness: Double;
CC topology: Unknown;
CC hypothetical: NO;
CC anti-sense: NO;
FH Key Location/Qualifiers
FT source 1..1704
FT mat_peptide 1..1704
FEATURES
source Location/Qualifiers
1..1704
/organism="Luciola lateralis"
/db_xref="taxon:7052"
BASE COUNT 541 a 275 c 361 g 527 t
ORIGIN
alignment_scores:
Quality: 2801.00 Length: 546
Ratio: 5.130 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.451
alignment_block:
US-09-581-241-4 x E12278 ..
Align seg 1/1 to: E12278 from: 1 to: 1704
3 AsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPheTyr 19
67 AGCTCTGACGACGATGAAATATTGTATGTCTGACCATTTTAAACC 116
19 oileglugluglyseralaglyalaglualeuarglytyrMetAspArg 36
117 TATTGAAGAGGATCTGCTGGAGCACATTCGCCAGATATATGATGAT 166
36 yralalysleuglyalalalealapherThrasnalaleuthrglyValasp 52
|||||

```

```

167 ATGCAAAACTTGAGCAATTGCTTTACTACCGCACTTACCGGTGCGAT 216
53 TyrThrTyrAlaGluTyrLeuGluIuysSerCysCysLeuGlyValale 69
217 TATAGCTACGCCGCAATACTTAGAAAAATCATCTGCTAGAGAGGCTT 266
69 ulysasnTyrglyLeuValValaspGlyArgIleAlaLeuCysSerGlu 86
267 AAGCAATTATGCTTGGTTGGTTGATGAGAGAAATTCGTTATCCAGTGA 316
86 snCysGluGluPhePheIleProValleuAlaGlyLeuPheIleGlyVal 102
317 ACTGTGAAGAGTCTTATTCCTGATTAAGCCGGTTATTTATAGCTGTC 366
103 GlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuValHis 119
367 GGTGGCTCCCACTAATAGATTACACTACCTGAAATGGTTGCACAG 416
119 rleuGlyIleSerIysProThrIleValPheSerSerIysIysGlyLeu 136
417 TTTAGGCATCTCTAAGCCACACATTTGATTTAGTTCTAAAAAGGATTA 466
136 sPlysValIleThrValGlnIysThrValThrAlaIleLysThrIleVal 152
467 ATAAAGTTATTAAGTGTACAAAAACGTAAGTACTATTAACCACTTGTT 516
153 IleuAspSerIysValAspTyrArgGlyTyrGlnSerMetAspAsp 169
517 ATATTGACACCAAGGTGATTTATAGAGTTATCATTCATGACGACACTT 566
169 eileLysLysAsnThrProGlnIysPheLysGlySerSerPheLysThr 186
567 TATTAAAAAAACACTCCACAGGTTTCAAGATCAAGTTTAAACATG 616
186 alGluValAsnArgLysGluGluValAlaLeuIleMetAsnSerSerGly 202
617 TAGAAGTTAACGCCAAAGAACAAAGTGTCTTATTAATGAACCTTCGCGGT 666
203 SerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLeuVal 219
667 TCACCGGTTTGCCAAAAGGTGTCACACTTCTCATGAAATTTGGTTCAC 716
219 rArgPheSerHisAlaArgAspProIleTyrGlyAsnGluValSerPro 236
717 GCGTTTTTCTCACGCTAGAGATCCAAATTATGAAACCAAGTTTCACAG 766
236 lYThrAlaIleLeuThrValValProPheHisHisGlyPheGlyMetPhe 252
767 GCACGGCTATTTTACGTGTAGTACCATTCACATGCGTGTGGTATGTT 816
253 ThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLeuThr 269
817 ACTACTTTAGGCTATCTAAGTGTGGTTTCGTATTCATGATTAAGAA 866
269 sPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysCysSer 286
867 ATTTGACGAAGAGACTTTTAAAAACACGCAAGATTACAAATGTTCAA 916
286 erValIleLeuValProThrLeuPheAlaIleLeuAsnArgSerGluLeu 302
917 GCGTATTTCTGTACCGACCTTGTGTCATTTCTTAATACAGTAATTA 966
303 LeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGlyVal 319
967 CTCGATAAATATATTTATCAAAATTAGTGAATGTCATCTGGCGAGC 1016
319 aProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheAsnLeu 336
1017 ACCTTATCTTAAAGAAATGGTGAACCTGTTCTAAGCTTTAAATTTAC 1066
336 roGlyValArgGlnGlyTyrGlyLeuThrGlnThrHisSerAlaIle 352
1067 CCGGTGCTGCTCAAGGCTATGTTTAAACAGAAACCACTCTGCAATTAT 1116

```

```

353 11ethrProgluGlyAspAspLysProGlyAlaSerGlyLysValValPro 369
|||||
1117 ATCACACCGGAGGCGATGATTAACAGAGTGCTGCTGCGCAAGATTGTC 1166
369 OleuHeuLysAlaLysValIleAspLeuAspThrLysThrLeuGlyP 386
|||||
1167 ATTATTATTAAGCAAAAGTTATCGATCTTGATCTAAATAAACTTTGGGCC 1216
386 roAsnArqrGlyGlyLysValCysValLysGlyProMetLeuMetLysGly 402
|||||
1217 CGAACGACGCTGAGAGATTGTTGTAAGGGTCCATGCTTATGAAGGT 1266
403 TyrValAspAsnProGluAlaThrArgLysIleLeuAspGluGlyTyr 419
|||||
1267 TATGTGATTAATCCAGAACCAAGCAAGAAATCATGATGAAGAGGTGG 1316
419 pleuHisThrGlyAspIleGlyTyrTyrAspGlyGlyLysHisPhePhe 436
|||||
1317 GTTGCACACGAGGATATTGGGATTACGATCAAGAAAAACTTTCTTTA 1366
436 1LeValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGluValPro 452
|||||
1367 TCGTGATGCTTTGAAGCTCTTAATCAATACAAAGATATCAAGTACCA 1416
453 ProAlaGluLeuGlySerValLeuLeuGlnHisProAsnIlePheAspAl 469
|||||
1417 CCTGCTGAATTAAGAACTGTTCTTTTGAACATCCAAATATTTTGTATCC 1466
469 aglyValAlaGlyValProAspProIleAlaGlyLysLeuProGlyAlaVal 486
|||||
1467 CGCGCTGCTGGCGTCCAGATCCTTACTGCTGAGCTTCCGGGAGCTGG 1516
486 alValValLeuLysLysGlyLysSerMetThrGluLysGluValMetAsp 502
|||||
1517 TTGTGTACTTGAAGAAAGCAAAATCTATGCTGCAAAAGAAAGTAAATGAT 1566
503 TyrValAlaSerGlyValSerAsnAlaLysArgLeuArgLysIleValAla 519
|||||
1567 TACGTGCTAGTCAAGTTTCAAAATGCAAAACGTTTGGTGCTGCTCCG 1616
519 gPheValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLys 536
|||||
1617 TTTTGTGAGAGGATACCTTAAGGTCTCACTGCTAAATTAAGACGTAAAG 1666
536 1aIleArgGluIleLeuLysLysProValAlaLysMet 548
|||||
1667 CAATTAGAGAAATACGTGAAGAAACAGACTGCTAAGATG 1704
seq_name: gb_pat: E13411
seq_documentation_block: 1704 bp DNA PAT 24-JUN-1998
LOCUS E13411
DEFINITION cDNA encoding biotinylated firefly luciferase bl203.
ACCESSION E13411
VERSION E13411.1 GI:3252216
KEYWORDS JP 1997187281-A/2.
SOURCE unidentifed.
ORGANISM unidentifed.
REFERENCE 1 (bases 1 to 1704)
AUTHORS Masuda,S., Tatsumi,H. and Koyama,T.
TITLE ANTIBODY-FIREFLY LUCIFERASE FUSION PROTEIN, GENE OBTAINED THEREFROM, NOVEL RECOMBINANT DNA AND PRODUCTION OF ANTIBODY-FIREFLY LUCIFERASE FUSION PROTEIN
JOURNAL Patent: JP 1997187281-A 2 22-JUL-1997;
COMMENT KIKKOMAN CORP
OS Luciola lateralis (firefly)
PN JP 1997187281-A/2
PD 22-JUL-1997
PR 09-JAN-1996 JP 1996001812
PI MASUDA SUSUMU, TATSUMI HIROKI, KOYAMA TAJIJI
PC C12N15/09,A61K39/395,C07H21/04,C07K14/47,C07K19/00,C12P21/02,

```

```

PC G01N33/535//
PC C12N9/02,(C12P21/02,C12R1:19);
CC strandedness: Double;
CC topology: Unknown;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH source 1..1704 /organism='Luciola lateralis' FT
FT /clone='pHLf200' /product='Biotinylated firefly luciferase FT
FT mat_peptide 1..1704
FT /db_xref='taxon:32644'
FEATURES
source location/Qualifiers
1..1704 /organism='unidentifed'
/db_xref='taxon:32644'
BASE COUNT 541 a 275 c 361 g 527 t
ORIGIN
alignment_scores:
Quality: 2801.00 Length: 546
Ratio: 5.130 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.451
alignment_block:
US-09-581-241-4 x E13411 ..
Align seg 1/1 to: E13411 from: 1 to: 1704
3 AsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPheTyrPr 19
|||||
67 AGTCTCGAAGACGATGAAATATGTTGATGCTCGAAGCATTTACCC 116
19 oIleGluGlySerAlaGlyAlaGlnLeuArgLysTyrMetAspArgT 36
|||||
117 TATTAAGAGGGGATCTGCTGGAGCAATTCGGCAAGATATGATGAT 166
36 TyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGlyValAsp 52
|||||
167 ATGCNAACCTTGAGCAATGCTTTACTACGCACTTACCGGTGTGAT 216
53 TyrThrTyrAlaGluTyrLeuGluLysSerCysLeuGlyGluAlaLe 69
|||||
217 TATAGTACGCCGCAATTAAGTAAATCATGCTGTAGAGAGGCTTT 266
69 uLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuLysSerGlu 86
|||||
267 AAAGAAATTAATGTTGGTTGGTGTGATGAGAGAAATTCGTTATGACAGTAA 316
86 snCysGluGluPhePheIleProValIleuAlaGlyLeuPheIleGlyVal 102
|||||
317 ACTGTGAAGAGTTCTTATCTCTGATTTAGCGGTTTATTTATAGTGTG 366
103 GlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuValHis 119
|||||
367 GGTGTGCTCCAACTAATGAGATTACACTACGTAAGTAATGCTTCACAG 416
119 rLeuGlyIleSerLysProThrIleValPheSerSerLysGlyLeuAla 136
|||||
417 TTTTGGCATCTCTTAACCAACATTTATTTAGTTTAAAGAGATTAG 466
136 sPlyValIleThrValGlnLysThrValThrAlaIleLysThrIleVal 152
|||||
467 ATTAAGTTTAATCTATACAAAAAGCTAAGCTATTAACCATTTGTT 516
153 IleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAspAsn 169
|||||
517 ATATTGACAGCAAGATGATTAATAGAGTTATCAATCATGACAACTT 566
169 eIleLysLysAsnThrProGlnIlePheLysGlySerSerPheLysThrV 186

```

```

|||||
567 TATTAATAAAACACTCCACAAAGTTTCAAGATCAAGTTTAAACAG 616
186 aIglValaIasnArqLysgluValaIaIeulIleMetasnSerSergly 202
617 TAGAAGTTAACCGCAAGAACAAAGTTGCTTATATAGAACCTTTCGGGT 666
203 SerThrIgluLeuProLysgluValaIeulThrHisgluAsnLeuValThr 219
667 TCAACCGGTTTCCAAAAGGTGTGCAACTTACTCATGAATAATTGGTAC 716
219 rArpHeserHisAlaArgAspProIlePyrGlyAsnGluValSerProG 236
717 GCGTTTCTCAGCTAGAGATCCAAATTATGGAACCAAGTTTACCAG 766
236 lYThrAlaIleLeuThrValaIProPheHisIsglyPheGlyMetPhe 252
767 GCACGGCTATTTTAACTGAGTACCATTCACATGCTTTTGATGTTT 816
253 ThrThrIleuGlyTyrLeuThrCysGlyPheArgIleValMetLeuThrly 269
817 ACTACTTTAGGCTATCTACTTGTGTTTGTATGTCATGTTAAACGAA 866
269 sPheAspGluGluThrPheLeuLysThrLeuGluAspTyrLysCysSers 286
867 ATTGACGAAAGAACTTTTAAAAACACCTGCACAGATTCAAAATGTTCA 916
286 eValaIleLeuValProThrLeuPheAlaIleLeuAsnArgSergluLeu 302
917 GCGTTATTTCTGTACCGACTTTGTTGCAATTCTTAATAGAAAGTCAATTA 966
303 LeuAspLysTyrAspLeuSerAsnLeuValaIleuIleAsnSerglyVal 319
967 CTCGATTAATATGATTTATCAAAATTTACTTGAATTTGCATCTGGCGAGC 1016
319 aProLeuSerLysgluIleGlyAlaValaIaIaArgArpPheAsnLeuP 336
1017 ACCTTTATCTAAAGAAATGGTGAAGCTGTTGTCAGACGTTTAAATTAC 1066
336 rGluValaIaIaGlnGlyTyrGlyLeuThrGluThrThrSerAlaIleIle 352
1067 CCGGTGTTCGCAAGCTATGTTTAAACAGAAACACCTCGCAATTAAT 1116
353 IleThrProgluLysAspLysProgluAlaSerGlyLysValaIaIaPr 369
1117 ATCAGACCCGGAAGCGATGATAAACGAGTGTCTTGCAAAAGTTGTGCC 1166
369 oLeuPheLysAlaLysValaIleAspLeuAspThrLysLysThrLeuGlyP 386
1167 ATTAATTTAAAGCAAAAGTTATCGATCTTGATTAATAAAAAAATTTGGGCC 1216
386 rAsnArqArqGlygluValaIaIaLysgluPheMetLeuMetLysGly 402
1217 CGAAGCAGACTGGAAGAAAGTTTGTGTAAAGGGTCTTATGCTTAAGAAGGT 1266
403 TyrValaIaIaAsnProgluAlaThrArgGluIleIleAspLugluGlyTr 419
1267 TATGTAGATATATCCAGAACCAACAGAAATCATATAGTAAGAAGTTG 1316
419 pleuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisPhePheT 436
1317 GTTGCCACACGAGATATGGGTATTAACGATGAAGAAAAAATTTCTTTTA 1366
436 lValaIaIaAspArgLysSerLeuIleLysTyrLysgluTyrGluValPro 452
1367 TCGTGTATGTTGAGTCTTATATCAAAATACAAAGATATCAAGTACCA 1416
453 ProAlaGluLeuGluSerValLeuLeuGluHisProAsnIlePheAspAl 469
1417 CCGTGCGAATTAAGAAATCTGTTTGGCAACATCAAAATATTTTGTATGC 1466
469 aGlyValaIaIaGlyValaIaProAspProIleAgluGluLeuProgluAla 486
|||||

```

```

1467 CGGCGTTGCTGGCGCTTCAGATCCTATACCTGCTGAGCTTCCGGGAGCTG 1516
486 aValaIleuLysLysgluLysSerMetThrGluLysgluValaIaIaAsp 502
1517 TTGTTGTTACTGAAAAAGGAAATCTATGATGAAAAAGAAATATATGAT 1566
503 TyrValaIaIaSerGlnValSerAsnAlaLysArgLeuArgGlyGlyVala 519
1567 TACGTTGCTAGTCAAGTTTCAATGCAAAACGTTTGCCTGGTGGTCCG 1616
519 gPheValaIaIaAspGluValaIaProLysgluLeuThrGlyLysIleAspGlyLysA 536
1617 TTTTGTGACGAAAGTAAAGCTTAAAGCTCAGCTGTAATAATGACGTTAAAG 1666
536 lAlaIaIaGluLeuLysLysProValaIaIaLysMet 548
1667 CAATTAGAGAAATACTGAAGAAACCAAGTGTGAAGATG 1704

seq_name: em_pat:E10169
seq_documentation_block:
ID E10169 standard: DNA: UNC: 2019 BP.
XX AC E10169;
XX AC E10169;
SV E10169.1
XX 08-OCT-1997 (Rel. 52, Created)
DT 02-SEP-2000 (Rel. 65, last updated, Version 2)
XX DE DNA encoding streptoavidin-firefly luciferase fusion protein.
XX JP 1995289264-A/1.
XX OS unidentified.
XX OC unidentified.
XX RN [1]
RP 1-2019
RA Tatsumi H., Fukuda M.;
RT "VARIANT TYPE STREPTAVIDIN GENE, FUSION PROTEIN GENE OF VARIANT TYPE
RT STREPTAVIDIN-FIREFLY LUCIFERASE, NEW RECOMBINANT AND PRODUCTION OF FUSION
RT PROTEIN OF VARIANT TYPE STREPTAVIDIN-FIREFLY LUCIFERASE";
RL Patent number JP1995289264-A/1, 07-NOV-1995.
XX KIKKOMAN CORP.
XX OS None
XX OC Artificial sequences.
XX PN JP 1995289264-A/1
XX PD 07-NOV-1995
XX PF 27-APR-1994 JP 1994090275
XX PI TANSUMI HIROKI, FUKUDA MASARU
XX PC C12N15/09,C07K14/36,C07K19/00,C12P21/02,G01N3/53,(C12P21/02,
XX PC C12R1.19);
XX CC strandedness: Double;
XX CC topology: linear;
XX CC hypothetical: No;
XX CC anti-sense: No;
XX CC FH key location/Qualifiers
XX CC FH source 1..2019
XX CC FT "organism="Artificial sequences"
XX CC FT CDS 1..2019
XX CC FT /product="streptoavidin-firefly luciferase
XX CC FT fusion protein"
XX FH key location/Qualifiers
XX FH source 1..2019
XX FT /db_xref="taxon:32644"
XX FT /organism="unidentified"
XX SQ
Sequence 2019 BP; 590 A; 412 C; 467 G; 550 T; 0 other.

```


OM of: US-09-581-241-6 to: GenEmbl.* out_format : pfs

Date: Nov 29, 2001 3:27 AM

About: Results were produced by the GenCore software, version 4.5.
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09581241/runat_28112001_175311_14864/app_query.fasta_1.1226
-DB=GenEmbl -QFMT=fastlap -SUFFIX=rg -GAP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOAPCL=0.000 -LOAPEXT=0.000 -GAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-GAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500 -DELDP=6.000
-DELXT=7.000 -START=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DCCALIGN=200 -THR_SCORE=pc -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=200000000 -USER=US09581241@cgn1.1.9518 -NCPU=6
-ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPHY
-WAIT -THREADS=1

Search information block:

Query: US-09-581-241-6
Query length: 548
Database: GenEmbl.*
Database sequences: 1472140
Database length: 341344837
Search time (sec): 3171.350000

Score_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
gb_pat:E36480	+ 2823.00	3610.58	7.7e-193	1644	E36480 Luciferase and method
gb_pat:E36479	+ 2821.00	3608.01	1.1e-192	1644	E36479 Luciferase and method
gb_pat:AR098258	+ 2815.00	3600.32	2.9e-192	1644	AR098258 Sequence 13 from pate
gb_pat:E02495	+ 2814.00	3599.04	3.4e-192	1644	E02495 CDNA encoding luciferase
gb_pat:E05448	+ 2814.00	3599.04	3.4e-192	1644	E05448 CDNA sequence of luciferase
gb_pat:AR043317	+ 2812.00	3594.94	5.7e-192	1908	AR043317 Sequence 8 from patent
gb_pat:AR062709	+ 2812.00	3594.94	5.7e-192	1908	AR062709 Sequence 8 from patent
gb_pat:E12279	+ 2812.00	3594.94	5.7e-192	1908	E12279 CDNA encoding biotinylated
gb_pat:E16288	+ 2812.00	3594.94	5.7e-192	1908	E16288 DNA encoding fusion pro
gb_pat:AR043316	+ 2799.00	3579.44	4.2e-191	1704	AR043316 Sequence 5 from patent
gb_pat:AR062708	+ 2799.00	3579.44	4.2e-191	1704	AR062708 Sequence 5 from patent
gb_pat:E13411	+ 2799.00	3579.44	4.2e-191	1704	E13411 CDNA encoding biotinylated
gb_pat:E10169	+ 2799.00	3577.69	5.3e-191	2019	E10169 DNA encoding streptococ
gb_pat:E10170	+ 2799.00	3577.69	5.3e-191	2019	E10170 DNA encoding mutated st
gb_pat:E13412	+ 2798.00	3574.78	7.6e-191	2364	E13412 CDNA encoding fused pro
gb_pat:AR108849	+ 2684.00	3432.41	6.5e-183	1644	AR108849 Sequence 3 from patent
gb_pat:AR151719	+ 2684.00	3432.41	6.5e-183	1644	AR151719 Sequence 3 from patent
gb_pat:E01857	+ 2684.00	3432.41	6.5e-183	1644	E01857 DNA encoding luciferase
gb_pat:E02267	+ 2684.00	3432.41	6.5e-183	1644	E02267 CDNA encoding Photinus
gb_pat:E03259	+ 2684.00	3432.41	6.5e-183	1644	E03259 CDNA encoding luciferase
gb_pat:E05447	+ 2684.00	3432.41	6.5e-183	1644	E05447 CDNA sequence of luciferase
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR098253	+ 2596.50	3309.70	4.4e-176	4573	AR098253 Sequence 8 from patent
gb_pat:AR098255	+ 2596.50	3309.70	4.4e-176	4573	AR098255 Sequence 9 from patent
gb_pat:AR					

```

84  erclunscysgluiphepelleprovalleuallaglyleuphelle 100
    |||||||
251  GTGAAACCTGGAAGATCTTATCTCTATAGCGGTTATTATATA 300
    |||||||
101  GLVVALGLVALAPROTHRANGULIETHTHLEUARGGLULEUVA 117
    |||||||
301  GGGTGGGTGGTCTCACTAATGAGATTACACTCTACGTGAATGGT 350
    |||||||
117  LHSerLeuGlyLLeSerLysProThrLLeValLpHeSerSerLysG 134
    |||||||
351  TCACAGTTAGGATCTCTAAGCCACAATGTGATTAGTTCAAAAAAG 400
    |||||||
134  LyeuaspLysValLLeThrValGluLysThrValThraLLeLysThr 150
    |||||||
401  GATTGATTAAGTTATATACGTACAAAAACGGTAACCTGCTATTAACC 450
    |||||||
151  LLeValLLeLeuaspSerLysValAspTYRArgLTYRGLInSerMetas 167
    |||||||
451  ATTGTTATATTGGACACGACAAAGTGATTTAGAGGTTATCAATCCATGA 500
    |||||||
167  pAsnPhelLLeLysLysAsnThrProGlnGlyPheLysGlySerSerPheL 184
    |||||||
501  CAACCTTATTAATAAAAAACACCTCCACAGGTTCAAGATCAAGTTTAA 550
    |||||||
184  YsthrValGluValAsnArgLysGluGlnValAlaLeuLLeMetasSer 200
    |||||||
551  AAcGTGTAGAGTTAAACCGCAAGAACAGGTTGCTCTTATTAAGAAGCTCT 600
    |||||||
201  SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnI 217
    |||||||
601  TCGGGTTCACCGGTTGGCAAAAGGTGTGCAACTTACATCAATAATAT 650
    |||||||
217  eValThraArgPheSerHisAlaArgAspProLLeTYRGLYasnGlnValS 234
    |||||||
651  CcTCACTAATTTTTCACAGCTAGAGATCCAAATTATGGAACCAAGTTT 700
    |||||||
234  ePProGlyThraLLeLeuThrValAlaProPheHisHisGlyPheGly 250
    |||||||
701  CACCGAGCCGCGCTATTTTAACTGATGACCATTCATCATGTTGGT 750
    |||||||
251  MetPheThrThraLeuGlyTYRLeuThrCysGlyPheArgLLeValMetLe 267
    |||||||
751  ATGTTTACTACTTTAGGCTATCTAAGCTTGTGTTGATTTGCTATGTT 800
    |||||||
267  uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTYRLYsc 284
    |||||||
801  AACGAAATTTGACGAGAGACACTTTTAAAAACACCTGCAAGATTCAAAAT 850
    |||||||
284  YSerSerValLLeLeuValProThrLeuPheAlaLLeLeuSnaArgSer 300
    |||||||
851  GTTCAAGGCTTATCTGTGACCGACTTTGTTGCAATTTCTTATAGAAAT 900
    |||||||
301  GluLeuLeuaspLysTYRAspLeuSerSnaLeuValGluLLeAlaSerG 317
    |||||||
901  GAATTACTCGATTAATGATTTATCAAAATTTAGTTGAATGCAATCTGG 950
    |||||||
317  YGLValAProLeuSerLysGluLLeGlyGluValAlaAlaArgArgPheA 334
    |||||||
951  CGGACGACCTTATCTAAAGAAATGGTGAACCTGTGCTAGACGCTTTTA 1000
    |||||||
334  snLeuProGlyValArgGlnGlyTYRGLYLeuThrGluThrSerAla 350
    |||||||
1001  ATTTACCGGGGTGTGTCAGGCTATGTTTAAACGAAACCACTCTGCA 1050
    |||||||
351  LLeLLeLLeThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
    |||||||
1051  ATTATTTATCACACCGAAGCGGATGATTAACCAAGGTGCTTGCAAAAGT 1100
    |||||||
367  LValProLeuPheLysAlaLysValLLeAspLeuSphrLysLysThrL 384
    |||||||
1101  TGTGCAATTTATTAAGCAAAAGTTATCGATCTTACTATAAAAAACCTT 1150
    |||||||
384  euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
    |||||||

```

```

1151  TGGGCCGCGACAGACGCTGAGAAATTTGTGTAAGAGGCTCTATGCTTATG 1200
    |||||||
401  LysGlyTYRValAspAsnProGluAlaThrArgGluLLeLLeAspGluG 417
    |||||||
1201  AAAGTTATCTAGTAAATCCAGAAGCAAGAAAGAAATCATAGATGAAGA 1250
    |||||||
417  uGlyTrpLeuHisThrGlyAspLLeGlyTYRArgLysGluLysHisP 434
    |||||||
1251  AGGTTGTTCCACACAGAGATATTTGGTATTTAGATGAAGAAACAACTT 1300
    |||||||
434  hePheLLeValAspArgLeuLysSerLeuLLeLysTYRGLYTYRGLn 450
    |||||||
1301  TCTTTATCGTGATTCGTTGACGCTTTATCAATTAACAAGATATACAA 1350
    |||||||
451  ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnLLeP 467
    |||||||
1351  GTACCACTGCTGTAATTAGAATCTGTTCTTTTCAACATCCAAATATTTT 1400
    |||||||
467  eAspAlaGlyValAlaGlyValProAspProLLeAlaGlyGluLeuProG 484
    |||||||
1401  TGATGCCGGGCTTGTGCGCTTCCAGATCTTATAGCTGGTGAGCTTCGG 1450
    |||||||
484  LValAlaValValLeuLysGlyLysSerMetThrGluGlyValAl 500
    |||||||
1451  GAGCTGTTGTTGACTTAAGAAAGAAATCTATGACTGAAAAAGAAATGA 1500
    |||||||
501  MetAspTYRValAlaSerGlnValSerAsnAlaLysArgLeuArgLyl 517
    |||||||
1501  ATGATTAGCTTCTGCTAGTCAAGTTTCAAAATGCAAAACGTTGGTGCTG 1550
    |||||||
517  YValArgPheValAspGluValProLysGlyLeuThrGlyLysLLeAspG 534
    |||||||
1551  TGTCCGTTTGTGAGAGAGTACCTTAAGGTCTCACTGGTAAATTTGAGG 1600
    |||||||
534  LYLysAlaLLeArgGluLLeLeuLysLysProValAlaLysMet 548
    |||||||
1601  GTAAGCAATTAGAGAAATAGCAAGAAACCAAGTTGCTTAAGATG 1644
    |||||||
seq_name: gb_pat:E36479
seq_documentation_block: 1644 bp DNA PAT 07-FEB-2001
LOCUS E36479
DEFINITION Luciferase and method for assaying intracellular ATP by using the
same.
ACCESSION E36479
VERSION E36479.1 GI:13022686
KEYWORDS JP 1999239493-A/3.
SOURCE Luciola lateralis.
ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
Elateriformia; Cantharoidae; Lampyridae; Luciola.
REFERENCE Noriaki,H.S.M.M.
AUTHORS
TITLE Luciferase and method for assaying intracellular ATP by using the
JOURNAL KIKKOMAN CORP
COMMENT OS Luciola lateralis
PN JP 1999239493-A/3
PD 07-SEP-1999
PF 21-DEC-1998 JP 1998363108
PR
PI NORIAKI HATTORI,SEIJI MURAKAMI
PC C12N15/09,C12N9/02,C12Q1/66//C12N9/02,C12R1:19,C12N15/00 CC
FEATURES
source location/Qualifiers
FT key CDS (1)..(1644).
/db_xref="taxon:7052"
BASE COUNT 528 a 261 c 349 g 506 t

```



```

1401 TGATCGCGCGCTGCTGCGCTTCACATCCTATAGCTGTGAGCTCCGG 1450
484 1yAlaValAlaValLeuLysGlyLysSerMetThrGluLysGluVal 500
1451 GAGCTGTGTGTACTTACGAAAGAAATCTATGACTGAAAGAAAGTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgLysG 517
1501 ATGATTACGTTGCTAGTCAGATTCAAAATGCAAAACGTTGCGTGTGG 1550
517 ValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
1551 TGTCGCTTTTGTGACGAGTAAAGTCTACGTCACGTAAATGTGACG 1600
534 1yLysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
1601 GTAAAGCAATTAGAGAAATGACTGAAAGAACCCAGTTGCTAAGATG 1644

seq_name: gb_Pat:E02495
seq_documentation_block:
LOCUS E02495 1644 bp RNA PAT 29-SEP-1997
DEFINITION CDNA encoding Luciferase.
ACCESSION E02495
VERSION E02495.1 GI:2170725
KEYWORDS JP 1990171189-A/1.
SOURCE
ORGANISM
Luciola lateralis.
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
Elateriformia; Cantharoidae; Lampyridae; Luciola.
REFERENCE
AUTHORS KajiYama,N.,Tatsumi,H. and Nakano,E.
TITLE Luciferase GENE
JOURNAL Patent: JP 1990171189-A 1 02-JUL-1990;
COMMENT KIRKMAN CORP
OS Luciola lateralis
PN JP 1990171189-A/1
PD 02-JUL-1990
PF 22-DEC-1988 JP 1988322029
PI KAJIYAMA NAKOI, TATSUMI HIROKI, NAKANO EIICHI PC
C12N15/53/C12N9/02;
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH key Location/Qualifiers
FT mat_peptide 1..1644
FT FT /product='Luciferase'.
FEATURES
source location/Qualifiers
1..1644
/organism='Luciola lateralis'
/Db_xref='taxon:7052'
BASE COUNT 529 a 262 c 349 g 504 t
ORIGIN
alignment_scores:
Quality: 2814.00 Length: 548
Ratio: 5.144 Gaps: 0
Percent Similarity: 99.818 Percent Identity: 99.635
alignment_block:
US-09-581-241-6 x E02495
Align seg 1/1 to: E02495 from: 1 to: 1644
1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17
1 ATGCAAAACATGAGAAACATGAAATATTTGTATAGTCTGACCAACCAT 50
17 eTyrProIleGluGluGlySerIleGlyAlaGlnLeuArgLysTyrMet 34

```

951 CGAGACACCTTTATCTAAGAAATGGTGAAGCTGTGCTAGACGTTTAA 1000
 334 subeupProgiValArgInglYTrGlyLeuThrgluThrThSerAla 350
 1001 ATTACCGGGTGTTCGCAAGGCTATGTTTACAGAAAACACTCTGCA 1050
 351 IleIleIleThrProgiUgiYAspAspLysProgiYAlaSerGlyLysVa 367
 1051 ATTATTATCACACCGAAGGCGATGATTAACCAAGGTCCTCTGGCAAGT 1100
 367 ValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
 1101 TGTGCCATTTATTAAGCAAAAGTTATCGATCTTGATCTAAAAAACTT 1150
 384 euGIyProAsnArgArgLysLysValLysValLysGlyProMetLeuMet 400
 1151 TGGGCCCCAGAACACGCGAGAAAGTTGTCTAAAGGTCCTATGCTTATG 1200
 401 LysGlyTyrValAspAsnProgiUalThrArgLysIleLeuAspGluG 417
 1201 AAAGGTTATGTAGATATATCCAGAACACAAAGAAATCATAGATGAGA 1250
 417 uGIYTrPLeuHisThrGlyAspIleGlyTyrTyrAspGluGlyLysHisP 434
 1251 AGGTTGGTTTCACACAGAGATATTGGGATTTACGATGAAAGAAAACATT 1300
 434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
 1301 TCTTTATCGTGATCGTTTGAAGCTTTTATCAAAATACAAGATATNCAA 1350
 451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
 1351 GTRACCACTGCTGAATTAAGATCTGTTCTTTCACATCCAAATATTTT 1400
 467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyLysLeuProG 484
 1401 TGATGCCGGCGCTTGGCGCTTCACATCTATAGCTGGAGACTTCCGG 1450
 484 ValAlaValAlaValLeuLysLysGlyLysSerMetThrgluLysGluVal 500
 1451 GAGCTGCTGTGTGACTTGAAGAAAAGAAATCTGACTGAAAAGAAAGTA 1500
 501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgLysG 517
 1501 ATGGATTACGTTGCTAGTCAGATTTCAATGCAAAACGTTTGGCTGTGG 1550
 517 ValAlaArgPheValAspGluValProLysGlyLeuThrglyLysIleAspG 534
 1551 TGTCCGTTTGTGGAGAGAGTACTTAAGGTCTCACTGTAATAATTGACG 1600
 534 LysAlaIleAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
 1601 GTAAAGCAATTTAGAGAAATCTGAAGAAACAGTTGCTAAGATG 1644

seq_name: gb_pat:E05448
 seq_documentation_block:
 LOCUS E05448 1644 bp RNA PAT 29-SEP-1997
 DEFINITION cDNA sequence of Luciferase.
 ACCESSION E05448
 VERSION E05448.1 GI:2173637
 KEYWORDS JP 1993244942-A/2.
 SOURCE JP 1993244942-A/2.
 ORGANISM Luciola lateralis.
 Luciola lateralis.
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Plekoryeta; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 Elateriformia; Cantharoidea; Lampyridae; Luciola.
 REFERENCE 1 (bases 1 to 1644)
 AUTHORS Kajiyama, N. and Nakano, E.
 TITLE HEAT-STABLE FIREFLY LUCIFERASE, HEAT-STABLE FIREFLY LUCIFERASE
 GENE, NEW RECOMBINANT DNA AND PRODUCTION OF HEAT-STABLE FIREFLY
 LUCIFERASE
 JOURNAL Patent: JP 1993244942-A 2 24-SEP-1993;

COMMENT KIRKMAN CORP
 OS Luciola lateralis
 PN JP 1993244942-A/2
 PD 24-SEP-1993
 PF 22-MAY-1992 JP 1992131057
 PR 27-JUN-1991 JP 91P 157117, 29-NOV-1991 JP 91P 317064 PI
 PC KAJIYAMA NAKO, NAKANO EIICHI
 PC C12N9/02, C12N15/53//C12N1/21, (C12N9/02, C12R1.19); CC
 strandness: Double;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 FH Key location/Qualifiers
 FT CDS 1..1644 /product='Luciola lateralis luciferase'.
 FEATURES source location/Qualifiers
 1..1644 /organism='Luciola lateralis'
 /db_xref='taxon:7052'.
 BASE COUNT 529 a 262 c 349 g 504 t
 ORIGIN

alignment_scores:
 Quality: 2814.00 Length: 548
 Ratio: 5.144 Gaps: 0
 Percent Similarity: 99.818 Percent Identity: 99.635

alignment_block:
 US-09-581-241-6 x E05448 ..
 Align seg 1/1 to: E05448 from: 1 to: 1644

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProh 17
 1 ATGGAAGAACTGTGAGAACGATGAAATATTTGTATGCTGTAACCAT 50
 17 eYrProIleGluGluGlySerAlaGlyValAlaGluLeuArgLysTyrMet 34
 51 TTTACCTATTGGAAGAGGATCTCTGAGACGACATTTGCCGACGATATG 100
 34 sPArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrgly 50
 101 ATCGATATGCAAAACCTTGAGCATTTGCTTTTACTTAACGACATTACCG 150
 51 ValAspTyrThrTyrAlaGluLysLysSerGlyLysIleAspG 67
 151 GTGGAATTATACGACCGCGAATCTTAGAAGAAATCATGCTGTCTAGAGA 200
 67 uAlaLeuLysAsnTyrGlyLeuValAlaAspGlyArgIleAlaLeuGly 84
 201 GGCTTAAAGAAATTTATGTTGGTGTGTGATGGAAGAAATTCGCTTATG 250
 84 eGIuAsnGlyGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
 251 GTGAAGAACTGTGAAGAAATCTTATCTCTGATTTACCGGTTATTATTA 300
 101 GlyValGlyValAlaProThrsngluIleTyrThrLeuArgLysLeu 117
 301 GGTGTGCTGTGCTCCAACTAATGATTTACACTCTACGTAATGTTG 350
 117 IhSerLeuGlyIleSerLysProThrIleValPheSerLysLysG 134
 351 TCACAGTTTAGGCATCTCTAAGCAACATTTGATTTACTTCAAAAAG 400
 134 LysLeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
 401 GATTAGATTAAGTTATTACTGTACAAAAACGTAACCTATTAAACC 450
 151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
 451 ATTGTTATTTGGACAGCAAGTGGATTTAGAGGTTATCAATCATAGTA 500

167 pasnphelllelylsasnThrProGlnGlyPheLysGlySerSerPheL 184
 1401 TGAATGCGGCGTGGCTGGCGCTTCCAGATCTTACTGCTGACCTCCGG 1450
 501 CAACTTATTAATAAAGACCTCCACAAAGGTTCAAAAGATCAAGTTTAA 550
 184 ysthrrValIglValasnArgLysGluGlnValAlaLeuIleMetanser 200
 551 AAAGTGAAGAGTTAAACGCAAGAACAAAGTCTCTTATTAAGAACCTCT 600
 201 SerGlySerThrGlyLeuProLysGlyValAlaGlnLeuThrHisGluasnI 217
 601 TCGGCTTCAACCGGTTTGCACAAAGGTGTCACACTTACTCATGAAATGC 650
 217 evalThrArgPheSerHisAlaArgaspProIleTyrGlyAsnGlnValS 234
 651 AGTCACAGATGATTTCTCCACGCTAGAGATTCCAATTTATGAAACCACTT 700
 234 erProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
 701 CACCAAGCGCGCTATTTTAACTGATGATACCATTCATCGTTTGGT 750
 251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetIle 267
 751 ATGTTTACTACTTATAGCTATCTAACTTGCGTTTCTGATATGTCATGTT 800
 267 uThrLysPheaspGluGlnuThrPheLeuLysThrLeuGlnAspTyrLysC 284
 801 AACCAAAATTTGACAGAGACTTTTAAAAACCTGCAAGATTAACAAT 850
 284 ysserSerValIleLeuValProThrPhePheAlaIleLeuAsnArgSer 300
 851 GTTCAAGCGTATTTCTTGTACCGACTTGTGTAATTCCTTAATAGAACT 900
 301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
 901 GAATTAATCGATTAATAATGATTATCAAAATTTAGTCAAAATTCATCTCG 950
 317 yGlyAlaProLeuSerLysGluIleGlyAlaValAlaArgArgPheA 334
 951 CGGAGCACCCTTATCTTAAGAAATTTGGTGAAGCTGTTCTGACAGCTTTTA 1000
 334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
 1001 AATTACGGGCTTGTGTCAGCATATGTTTAAACAAACCACTTCGCA 1050
 351 llellellethrProGlnGlyAspAspLysProGlyAlaSerGlyLysVa 367
 1051 ATTATTATCACACCGAAGCGCATGATTAACAGAGTCTTCTGCAAGAT 1100
 367 ValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
 1101 TGTGCAATTAATTAAGCAAAAGTTATGATCTTATCTAATAAAAACTT 1150
 384 euGlyProAsnArgArgLysGluValLysValLysGlyProMetLeuMet 400
 1151 TGGGCGCCGAGACGATGAGAGAGTTTGTAAAGGTCCTATGCTTATG 1200
 401 LysGlyTyrValAspAsnProGluAlaThrArgGluIlelleAspGlnG 417
 1201 AAAGGTTATGTAGATTAATCCAGAACCAAGAACAAATCATGATGAGAGA 1250
 417 uGlyTyrLeuHisThrGlyAspIleGlyTyrTyrAspGlnGluLysHisP 434
 1251 AGGTTGCTGTCACACAGAGATATGGGATTAACATGAGAGAAAAAATTT 1300
 434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
 1301 TCTTTATCGTGCATGCTTTGAAGCTTTTAAATCAAAATACAAAGATATCAA 1350
 451 ValProProAlaGluLeuGlnuSerValLeuLeuGlnHisProAsnIleP 467
 1351 GTACCAACCTGCTGAATTAAGATCTGTTCTTTGCAACATCCAAATATTTT 1400

467 easpAlaGlyValAlaGlyValProAspProIleAlaGlyLeuLeuProG 484
 1401 TGAATGCGGCGTGGCTGGCGCTTCCAGATCTTACTGCTGACCTCCGG 1450
 484 ValAlaValAlaLeuLysLysGlyLysSerMetThrGluLysGluVal 500
 1451 GAGCTGTTGTTGCTGCTGAAAAAGAAAAATCATGACTGAAAAAGAGTA 1500
 501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyG 517
 1501 ATGGATTACGTTGCTAGTCAAGTTTCAAAATGCAAAAGCTTGGCGTGG 1550
 517 yValArgPheValAspGlnValProLysGlyLeuThrGlyLysIleAspG 534
 1551 TGTCCGTTTGTGGAGAGAGTAAAGTCTCACTGCTAAATTTGACG 1600
 534 LysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
 1601 GTAAAGCAATTAGAGAAATAGAGAAACCAAGTCTGTAAGATG 1644

seq_name: gb_in:LLUCCI

seq_documentation_block:

LOCUS LLUCCI 1781 bp mRNA INV 24-NOV-1993
 DEFINITION L. lateralis luciferase mRNA.
 ACCESSION X66919.38389
 VERSION X66919.1 GI:9526
 KEYWORDS luciferase.
 SOURCE Japanese firefly.
 ORGANISM Luciola lateralis
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 Elateriformia; Cantharoidae; Lampyridae; Luciola.

REFERENCE
 1 (bases 1 to 1781)
 2 (bases 1 to 1781)
 1 Tatsumi, H., Kajiyama, N. and Nakano, E.
 Molecular cloning and expression in *Escherichia coli* of a cDNA
 clone encoding luciferase of a firefly, *Luciola lateralis*
 Biochim. Biophys. Acta 1131 (2), 161-165 (1992)

JOURNAL
 MEDLINE
 92305054
 REFERENCE
 2 (bases 1 to 1781)
 1 Tatsumi, H.
 Direct Submission
 Submitted (21-JUL-1992) H. Tatsumi, Research and Development
 Division, Kikkoman Corporation, 399 Noda, Noda City, Chiba 278,
 JAPAN

FEATURES
 source location/Qualifiers
 1. 1781
 /organism="Luciola lateralis"
 /db_xref="taxon:7052"
 59. 1705
 /codon_start=1
 /product="luciferase"
 /protein_id="CAA47358.1"
 /db_xref="GI:9526"
 /db_xref="SWISS-PROT:O01158"
 /translation="MEMENDENTIYGEPEPTPIEBESAGQAKRYDRYAKUATAP
 TNAITGVDTYTAELVLEKSCLEALKNGLVVDGIALVCSNCEEFIPVLAGLFIGV
 GVAIPNEITLLELVHSLGISKPTIVFSKSKDKVITVQKTAIAKIVLIDSKVY
 RGYSQDNFKIKNTPOGFKSGSEFTVEVNRKEQVALINSSGSLGPGVOLTENAV
 TRFSHARDPIYGNVNSPGTALITVVPFHNGGMFTLGLYLCGPRIVMLTKFDETEL
 KTLQDYKCSVLTLPVTLALINRSLDKYLSNVEYASGAPLSEKIGAVARRN
 LRGVROGVLGETTSALITITPEGDKPGASGVVPLFRAKYIDLTDKTKLGPNNRGVY
 CVKGPMLKMGVVDNDEATREIIDEQGLHTDIDGYDEKHFIVDRKLSLTKRGIO
 VPAELESVLLQHPNIFDAGVAGVDPYIAGELPGAVVLEKGSMTKEVMDVYASOV
 SNAKRLRGVREVDEVPKLTGKIDGKAIRELTKRPVAKM"

CDS

BASE COUNT 579 a 276 c 367 g 559 t
 ORIGIN
 alignment_scores:
 Quality: 2814.00 Length: 548
 Ratio: 5.144 Gaps: 0
 Percent Similarity: 99.818 Percent Identity: 99.635

alignment block:

US-09-581-241-6 x LLLUCI ..

Align seg 1/1 to: LLLUCI from: 1 to: 1781

```

1 MetGluAsnMetGluAsnAspGluAsnIleValuYrGlyProGluProph 17
  |||||||
59 ATGGAAAACATGGAGAACGATGAAATATGCTATGCTGCAACCATTT 108
  |||||||
17 eYrProIleGluGlySerIleGluGluGluGluGluGluGluGluGlu 34
  |||||||
109 TTAACCTATGAGAGAGGATCTGCTGGAGACATTCGCCAATATATGG 138
  |||||||
34 spArgYrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
  |||||||
159 ATCGATATGCAAAACTTGGAGCAATGCTTTTACTAACGCACCTTACCGGT 208
  |||||||
51 ValAspYrThrTyraIleGluYrLeuGluYrLeuSerCysCysLeuGly 67
  |||||||
209 GTGATATATACGACCGCAATCTTAGAAAATCATGCTGTCTAGAGAGA 258
  |||||||
67 uAlaLeuLysAsnTyrglyLeuValValAspGlyArgIleAlaLeuCys 84
  |||||||
259 GCCTTTAAAGAAATTATGTTGGTGTGTGATGAGACAATTCGCTTATGCA 308
  |||||||
84 eGluAsnGlyGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
  |||||||
309 GTGAAACGCTGAGAAATCTTATTCGTATTCGATACCGGTTATTTATAA 358
  |||||||
101 GluValGlyValAlaProThrAsnGluIleYrThrLeuArgGluLeuVal 117
  |||||||
359 GGTGTGCGGTGGCTCCACTATGATGATTTACCTACGTGCAATTTGGT 408
  |||||||
117 HisSerLeuGlyIleSerLysProThrIleValPheSerSerLysLysG 134
  |||||||
409 TCACAGTTAGGATCTCTAACCCAACTGATTTACTTCTTAAAAAAG 458
  |||||||
134 ILeuAspLysValIleThrValGluLysThrValThrAlaIleLysThr 150
  |||||||
459 GATTTGATTAAGTTATTAAGTACCAAAAACGTAATCTCTTTTAAACC 508
  |||||||
151 ILeuAlaIleLeuAspSerLysValAspYrArgIleYrGluSerMetAs 167
  |||||||
509 ATTGTTAATTTGGACAGCAAGTGTATAGAGTTATCAATCAATGAGA 558
  |||||||
167 PAsnPheIleLysAsnThrProGluGlyPheLysGlySerSerPheL 184
  |||||||
559 CAACCTTTATTAATAAAACACTCCACAAGGTTTCAAAAGGATCAAGTTTA 608
  |||||||
184 YrThrValGluValAsnArgLysGluGluValAlaLeuIleMetAsnSer 200
  |||||||
609 AAACGTGTAAGTTAAACCGCAAAAGAAAGTTCGCTTATATGAACTCT 658
  |||||||
201 SerGlySerThrGlyLeuProLysGlyValGluLeuThrHisGluAsnI 217
  |||||||
659 TCGGGTTACCGGTTTCCAAAAGGTGTGCAACTTACTCATGAAATAAGCC 708
  |||||||
217 eValYrThrArgPheSerHisAlaArgAspProIleYrGlyAsnGluVal 234
  |||||||
709 AGTCACTGATTTTCTCACGCTAGAGATCCAAATTTATGGAACCAAGTTT 758
  |||||||
234 ePProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
  |||||||
759 CACCAGGACGCGTATTTTAACTGTAAGTACCATTCATCATGCTTGGT 808
  |||||||
251 MetPheThrThrLeuGlyYrLeuThrCysGlyPheArgIleValMetLe 267
  |||||||
809 ATGTTTACTACTTAGGCTATCTAACTTGCTGTTTCTGATTTGCATGTT 858
  |||||||
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGluAspYrLysC 284
  |||||||
859 AAGCAATTTGACGAGAGACTTTTAAAAAACACTGCAAGATTACAAAT 908
  |||||||

```

```

284 YrSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
  |||||||
909 GTTCAACCGTATTTCTGTACCGACTTTGTTGCAATTTCTTAATAGAGT 958
  |||||||
301 GluLeuLeuAspLysTyraPheSerAsnLeuValGluIleAlaSerG 317
  |||||||
959 GAATTTACTGATTAATGATTTATCAAAATTTAGTTGAATTCATCTCG 1008
  |||||||
317 YGlyAlaProLeuSerLysGluIleGlyAlaValAlaAlaArgPheA 334
  |||||||
1009 CGGAGCACCTTTATCTTAAAGAAATGGTGAAGCTGTGCTAGACGTTTA 1058
  |||||||
334 snLeuProGlyValArgGluGlyYrGlyLeuThrGluThrThrSerAla 350
  |||||||
1059 ATTTACCGGCTGTTCGCAAGGCTATGTTTAAACAAACACCTCTGCA 1108
  |||||||
351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
  |||||||
1109 ATTATTTATCACACCGGAGCGATGATTAACCAAGCTGCTTCTGCAAGT 1158
  |||||||
367 ValProLeuPheLysAlaLysValIleAspLeuAspThrLysThrL 384
  |||||||
1159 TGTGCCATTTTAAAGCAAAAGTTATCGATCTTGATCTTAAATAAACTT 1208
  |||||||
384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
  |||||||
1209 TGGGCCCGCAACAGCTGAGAGAGTTTGTGTAAGGCTCCTATGCTTAG 1258
  |||||||
401 LysGlyYrValAspAsnProGluAlaThrArgGluIleLeuAspGlu 417
  |||||||
1259 AAGGTTATGATGATATATCCAGAGCAACAGAAATATATATATCAAA 1308
  |||||||
417 uGlyTrpLeuHisThrGlyAspIleGlyYrTyraPrgGluGluYrHisP 434
  |||||||
1309 AGGTTGTTGCAACAGAGATATGGTATTAAGTGAAGAAAACATT 1358
  |||||||
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrglyGlu 450
  |||||||
1359 TCTTTATCTGATCGTTTGAAGCTTTTAAATCAAAATACAAAGATATCAA 1408
  |||||||
451 ValProProAlaGluLeuGluSerValLeuLeuGluHisProAsnIleP 467
  |||||||
1409 GTACACCGCTCGAATTAATCTGCTTTTGCACATCCAAATATTTT 1458
  |||||||
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
  |||||||
1459 TGATGCCGCGTGTGCGCTCCAGATCTTATAGCTGTGAGCTTCCGG 1508
  |||||||
484 IValAlaValAlaLeuLysLysGlyLysSerMetThrGluLysGluVal 500
  |||||||
1509 GAGCTGTTGTTTACTTGAAAAGAAATCTATGACTGAAAAAAGATA 1558
  |||||||
501 MetAspYrValAlaSerGluValSerAsnAlaLysArgLeuArgGly 517
  |||||||
1559 ATGATATAGCTGTGTAAGCAAGTTTCAATGCAAAACGTTCCGAGTGG 1608
  |||||||
517 YValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
  |||||||
1609 TGTCCGTTTGTGGACGAAGTACCTTAAAGCTCTCACTGTAATAATTTGACG 1658
  |||||||
534 IYLyAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
  |||||||
1659 GTAAAGCAATTTAGAAATACTGAAAGAACCGAGTTGCTAGAGATG 1702
  |||||||

```

seq_name: gb_pat:AR043317

seq_documentation_block:

LOCUS AR043317 1908 bp DNA

DEFINITION Sequence 8 from patent US 5614465.

ACCESSION AR043317

VERSION AR043317.1 GI:5964325

KEYWORDS

PAT 29-SEP-1999

501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyG1 517
 1501 ATGATTACCTGCTGCTAGTCAAGTTTCAATGCAAAACGTTGGCTGGTGG 1550
 517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
 1551 TGTGCGTTTGTGGACGAATGCTAAAGGTCTCAGCTGTAATAATTGACG 1600
 534 yLysAlaIleArgGluIleLeuLysLysProValAlaLys 547
 1601 GTAAGCAATTAAGAAATTAATCTAAGAAACCACTGCTTAAG 1641

seq_name: gb_pat:AR062709

seq_documentation_block:
 LOCUS AR062709 1908 bp DNA PAT 29-SEP-1999
 DEFINITION Sequence 8 from patent US 5843746.
 ACCESSION AR062709
 VERSION AR062709.1 GI:5990400
 KEYWORDS
 SOURCE
 ORGANISM
 UNKNOWN.
 UNCLASSIFIED.
 REFERENCE 1 (bases 1 to 1908)
 AUTHORS Tatsumi,H., Fukuda,S., Kikuchi,M. and Koyama,Y.
 TITLE Biotinlated firefly luciferase, a gene for biotinlated firefly
 luciferase, a recombinant DNA, a process for producing biotinlated
 luciferase and a bioluminescent analysis method
 JOURNAL Patent: US 5843746-A 8 01-DEC-1998;
 FEATURES
 SOURCE Location/Qualifiers
 1..1908
 /organism="unknown"
 BASE COUNT 598 a 329 c 426 g 555 t
 ORIGIN

alignment_scores:
 Quality: 2812.00 Length: 547
 Ratio: 5.141 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.634

alignment_block:
 US-09-581-241-6 x AR062709 ..

Align seg 1/1 to: AR062709 from: 1 to: 1908

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17
 1 ATGGAACATGAGAGACGATGAAATATGTGTATGTCCTGCAACCAT 50
 17 eTyrProIleGluGluGlySerAlaGlyAlaGluLeuArgLysTyrMet 34
 51 TTACCTTATGTAAGAGGATCTGTGGAGCACAAATGGCGCAAGTATATG 100
 34 sPArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
 101 ATCGATTACCAAACTTGAGCAATTCCTTTACTACGCACTTACCGGT 150
 51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGlyG1 67
 151 GTCGATTATACGTACGCCGAATTACTAGAAAAATCATGCTGTCTAGAGA 200
 67 uAlaLeuLysAsnTyrGlyLeuValAlaAspGlyArgIleAlaLeuCys 84
 201 GGCCTTTAAAGATTAATGTTGTGTGTGATGAGGAATGGCTATATGCA 250
 84 eArgLysCysGluGluPhePheIleProValIleuAlaGlyLeuPheIle 100
 251 GTGAAACTGTGAAGAGTCTTTATCTCTATTAAGCCGGTTATTTATA 300
 101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVa 117
 301 GGTGCGGTGTGGCTCCAACTAATGATTAACACTTACGTAATGTGT 350

117 hHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysLysG 134
 351 TCACAGTTTAGGCATCTCTAAGCCAAACAAATTTATTAAGTTTCAAAAAG 400
 134 yLeuAspLysValIleThrValGluLysThrValThrAlaIleLysThr 150
 401 GATTAGATAAAGTTATTAAGTACCAAAAAAGGTAAGTCTGCTTAAAGC 450
 151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
 451 ATTGTTATATTGGACAGCAAAAGTGAATATAGAGTTATCAATCAATGGA 500
 167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerPheL 184
 501 CAACCTTATTAATAAAACACCTCACAAAGGTTTCAAAAGATCAAGTTT 550
 184 yThrValGluValAsnArgLysGluIleValAlaLeuIleMetAsnSer 200
 551 AAACGTAGAGTTAACCGCAAAAGCAAGTCTCTTATTAATGAACCTT 600
 201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsn 217
 601 TCGGTTCAACCGTTTGGCAAAAGGTGCAACTTACATGAAATTT 650
 217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnVal 234
 651 GGTACACCGCTTTTCTACGCTAGAGATCCAAATTTAGAAACCAAGTT 700
 234 eProGlyThrAlaIleLeuThrValAlaProPheHisHisGlyPheGly 250
 701 CACACGACACGGCTATTTTACGTAGTACCATTCATCATGTTTGGT 750
 251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMet 267
 751 ATGTTTACTACTTACGCTATCTAATCTGTGTTTCCATTTGTCATGTT 800
 267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLys 284
 801 AACGAAATTTGACGAAGAGACCTTTTAAACACGCAAGTATCAAAAT 850
 284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
 851 GTTCAAGCGTTATCTTGTACCGACTTGTGCAATCTTAATAGAAAT 900
 301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSer 317
 901 GAATTACTGATAAATATGATTTATCAAAATTTAGTTGAATGCAATCTG 950
 317 yGlyAlaProLeuSerLysGluIleGlyLysAlaValAlaArgArgPhe 334
 951 CGGACACACTTATCTAAAGAAATGGTGAACCTGTGCTAGACGTTTA 1000
 334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrHisSer 350
 1001 ATTTACCGGGTGTGTCAAAGCTATGTTTAAACAGAAACCACTTCGCA 1050
 351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyVal 367
 1051 ATTATTATCACACCGGAAGCGCATATTAACCAAGTCTCTGCGCAAGT 1100
 367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysThrL 384
 1101 TGTGCATTAATTAAGCAAAAGTTATGATCTGATCTAATAAAACTT 1150
 384 eGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
 1151 TGGGCCCAAGACAGCTGGAAGAGTTGTGTAAAGGCTCATGTGCTTATG 1200
 401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGluG 417
 1201 AAAGTTATGTATGATTAATCCAGAGCAACAGAGAAATCATATGATGAAGA 1250

417 UGlyTrpLeuHisThrGlyAspIleGlyTyrTyrAspGluGlyHisP 434
 1251 AGGTTGGTGCACACAGAGATATTGGTATTAGATGAAGAAAAACATT 1300
 434 hepheileValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
 1301 TCTTTATCGTGATCGCTTTGAGCTTTATCAATACAAAGATATCA 1350
 451 ValProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
 1351 GTACCACCTGCTGAATTAGATCTGTTCTTTGCACATCAATATATT 1400
 467 eaSPALGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
 1401 TGATGCGGCGCTGCTGCGCTGCAGATCCATAGCTGATGAGCTTCGG 1450
 484 lYAlaValAlaValLeuLysGlyLysSerMetThrGlyLysGluVal 500
 1451 GAGCTGTTGTTGACTTGAAAAAGAAATCTATGACTGAAAAAGAGTA 1500
 501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
 1501 ATGATTCAGTTCCTAGTCAAGTTCAATGCAAAACGTTGCTGCTGG 1550
 517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAsp 534
 1551 TGTCGGTTTGTGACGAGTACTAAAGGCTCTCAGTGTAAATTAGC 1600
 534 lYlYSAlaIleArgGluIleLeuLysLysProValAlaLys 547
 1601 GTAAGCAATTAGAGAAATFACTGAAGAACCAAGTTGCTAAG 1641

seq_name: gb_pat:E12279

seq_documentation_block:

LOCUS E12279 1908 bp DNA PAT 24-JUN-1998
 DEFINITION cDNA encoding biotinylated Luciola luciferase.
 ACCESSION E12279
 VERSION E12279.1 GI:3251113
 KEYWORDS JP 1996308578-A/4.
 SOURCE Luciola lateralis.
 ORGANISM Luciola lateralis.
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 Elateriformia; Cantharoidae; Lampyridae; Luciola.
 1 (bases 1 to 1908)
 Tatum, H., Fukuda, M., Kikuchi, M. and Koyama, T.
 BIOTIN FIRE FLY LUCIFERASE, BIOTIN FIRE FLY LUCIFERASE GENE, NEW
 RECOMBINANT DNA, PRODUCTION OF BIOTIN FIRE FLY LUCIFERASE AND
 BIOLUMINESCENCE ANALYSIS
 Patent: JP 1996308578-A 4 26-NOV-1996;
 KIKKOMAN CORP
 OS Luciola lateralis
 PN JP 1996308578-A/4
 PD 26-NOV-1996
 PR 24-APR-1995 JP 1995098857
 PR 27-JUL-1994 JP 94P 193798, 14-MAR-1995 JP 95P 54625 PI
 TATSUMI HIROKI, FUKUDA MASARU, KIKUCHI MANORU, KOYAMA TAIJI PC
 C12N15/09,C12N9/02,C12O1/26,(C12N9/02,C12N1:185); CC
 strandedness: Double;
 CC topology: Unknown;
 CC hypothetical: No;
 CC anti-sense: No;
 FH key Location/Qualifiers
 FT source 1. 1908
 FT /clone='pHLf248'
 FT mat_peptide 1. 1908
 FT /product='biotinylated Luciola luciferase'.
 FT Location/Qualifiers
 1. 1908
 /organism='Luciola lateralis'

BASE COUNT 598 a 329 c 426 g 555 t
 ORIGIN

alignment_scores:
 Quality: 2812.00 Length: 547
 Ratio: 5.141 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.634

alignment_block:
 US-09-581-241-6 x E12279 ..

Align seg 1/1 to: E12279 from: 1 to: 1908

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17
 1 ATCGAAACATGAGCAAGCATGAAATATTGTATGTATGCTGGAACCAT 50
 17 eYrProIleGluGlySerAlaGlyAlaGluLeuAArgLysTyrMet 34
 51 TTACCTATTGAAAGAGGATCTCTGAGCACATTCGCCAGTATATG 100
 34 sPArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
 101 ATCGATATGCAAAACCTTGAGCAATTGCTTTACTAACGCACATTACCGGT 150
 51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysLeuGly 67
 151 GTCGATATTATGCTAGCCCGCATCTTAGAAATAATCATGCTGTCTAGGACA 200
 67 uAlaLeuLysAsnTyrGlyLeuValAlaAspGlyArgGlyAlaLeuCys 84
 201 GCCTTTAAAGATATTATGTTGTTGTTGATGAGAAATTCGCTTATGCA 250
 84 eRGLuAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPhe 100
 251 GTGAAACCTGTGAAGAGTCTTATTCTGTTATACCGGTTATTATTA 300
 101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeu 117
 301 GGTCGTGCTGCTCCACCTAATGAGATTATACCTTACGTAATTTGGT 350
 117 lHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysLys 134
 351 TCACAGTTAGGATCTCTAAGCAACATTTAGTTCTTAAAG 400
 134 lYLeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
 401 GATTAGATAAAGTTATTAAGTACAAAAAACGGTAACGCTATTATAAAC 450
 151 lIleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
 451 ATTGTTATATTGGACACCAAGTGATTTATAGAGTTATCAATCCATGGA 500
 167 pAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPhe 184
 501 CAACCTATTATTAATAAAACACTCCACAAGTTTCAAGAGTCAAGTTT 550
 184 yStrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
 551 AAACGTAGAGAGTTAACCGCAAGAACAAAGTGTCTTATAAAGAACTCT 600
 201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsn 217
 601 TCGGCTTCAACCGGTTTGGCAAAAGGTGTGCACTTCTCATGAAATTT 650
 217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnVal 234
 651 GGTCAAGCGGTTTCTCAGCTAGAGATCAATTTATGAAACCAAGTTT 700
 234 eRrProGlyThrAlaIleLeuThrValAlaProPheHisGlyPheGly 250

```

701 CACAGGACGCGCTATTCTAGTGTACCATTCATCATGTTTGGT 750
251 MetPheThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetIle 267
751 ATGTTTACTACTTAGGCTATCTTAAGCTTGGTTTGGTATGTCATGTT 800
267 uThrLysPheAspGluuThrPheLeuLysThrLeuGlnAspTyrLysC 284
801 AACGAATTTGACGAGAGACTTTTAAACACCTGCACAGATTACAAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTTCAAGCGTTATCTGTACCGACTTGTTCATTTCTTAATAGAACT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleLeuAsp 317
901 GATTTCTGATTAATATGATTTATCAAAATTTAGTTCAAAATTCATCTGG 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
951 CGAGACGACCTTATCTAAAGAAATGCTGAAGCTGTGCTAGACGTTTAA 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
1001 ATTTACCGGGTGTTCATCAAGGCTATGTTTAAACAGAAACACCTCTGCA 1050
351 IleIleIleThrProGluuLysAspAspLysProGlyAlaSerGlyLysVa 367
1051 ATTATATATACACCGGAGCGGATGATTAACAGAGTCTCTGCAAGCT 1100
367 yValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
1101 TGGGCGATTATTTAAAGCAAAAGTTATCATCTGTACTTAAAAAAAAGCT 1150
384 euGlyProAsnArgArgGlyGluValAlaCysValLysGlyPrometLeuMet 400
1151 TGGGCGCGAAGACAGCTGGAGAAAGTTGTGTAAGAGCTCTATAGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleLeuAspGlu 417
1201 AAAGTTAGTATGATTAATCAGAAAGCAACAGAAATCATGATGATGATG 1250
417 uGlyTyrLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
1251 AGGTGGTGTGCACACAGAGATTTGGGTAATTACGATGAAGAAAAACATT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
1301 TCTTATATCGTGGATCGTTGAAGTCTTTAATCAAAATACAAAGATATCA 1350
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
1351 GTTACACCGCTGAATTAGAATCTGTTCTTTGCAACATCCAAATATTTT 1400
467 eaSpAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
1401 TGATCGCGGCGGTGCTGCGCTCCAGATCTTATAGCTGCTGAGCTTCCGG 1450
484 yAlaValValValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
1451 GAGCTGTGTGTACTTGAAAAAGAAAAATCTATGACTGAAAAAGAGTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyL 517
1501 ATGGATTACGTTGCTAGTCAAGTTTCAAAATGCAAAACGTTTGGCGTGG 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
1551 TGTCTGTTTGTGGACGAAGTACCTAAAGGCTCTCACTGTTAAATTTGACG 1600
534 yLysAlaIleArgGluIleLeuLysLysProValAlaLys 547
1601 GTAAAGCAATTAGAGAAATACTGAGAAACCGAGTGTCTAAG 1641

```

```

seq_name: gb_pat:E16288
seq_documentation_block:
LOCUS E16288 1920 bp DNA PAT 28-JUL-1999
DEFINITION DNA encoding fusion protein which comprises luciferase and
lysoctaphin.
ACCESSION E16288
VERSION E16288.1 GI:5710971
KEYWORDS JP 1998150991-A/1.
SOURCE unclassified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1920)
AUTHORS Tatsumi,H., Fukuda,M. and Nagahara,A.
TITLE LUCIFERASE-LYSOCTAPHIN FUSED PROTEIN, ITS PRODUCTION AND
BIOLOGICAL ANALYSIS USING THE SAME
JOURNAL patent: JP 1998150991-A 1 09-JUN-1998;
KIKKOMAN CORP
COMMENT
OS None
OC Artificial sequences.
PN JP 1998150991-A/1
PD 09-JUN-1998
PF 25-NOV-1996 JP 1996328042
PI TATSUMI HIROKI, FUKUDA MASARU, NAGAHARA AYUMI PC
C12N15/09,C07H21/04,C07K19/00,C12N9/02,C12N9/52,C12P21/02, PC
C12Q1/66
PC (C12N9/02,C12R1:19),(C12N9/52,C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1..1920 /organism="Artificial sequences".
FT location/Qualifiers
1..1920
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 627 a 310 c 407 g 576 t
ORIGIN
alignment_scores:
Quality: 2812.00 Length: 547
Ratio: 5.141 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.634
alignment_block:
US-09-581-241-6 x E16288 ..
Align seg 1/1 to: E16288 from: 1 to: 1920
1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProh 17
|||||
1 ATGGAAGACATGAGAGACGATGATAATATTGTTAGTGTGCTGAACCAT 50
17 eTyrProIleGluGluGlySerAlaGlyValAlaGlnLeuArgLysTyrMet 34
|||||
51 TTACCTTATTGGAAGAGGATCTGCTGAGCACAAATTTGCCAAGTATATG 100
34 spArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
|||||
101 ATCGATTATCAAAACTTGAGCAATTTGCTTTACTAACGCACTTACCGGT 150
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysGlyLeuGly 67
|||||
151 GTCGATTATACCTACCGCGAATATCTTAGAAAAATCATGCTGTCTAGAGA 200
67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
|||||
201 GCGTTTAAAGATTTATGTTTGTGTTGATGAGAAATTCGCTTATATGCA 250

```

```

84  erGluAsnGlyGluPhePheIleProValLeuAlaGlyLeuPheIle 100
|||||
251  GTGAAACCTGGAAGAGTCTTATTCCTGATATAGCCGGTTATTTATA 300
|||||
101  GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVa 117
|||||
301  GGTGTGGGTGGCTCCAACTAATGAGATTTACCTGACGTGATTTGGT 350
|||||
117  LHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysG 134
|||||
351  TCACAGTTTAGCATCTCTAAGCCACAACTGTTAGTTCTTAAAAAG 400
|||||
134  LLeuAspLysValIleThrValAlaGlnThrValIleThrAlaIleLysThr 150
|||||
401  GATTAGATTAAGTTATTAAGTACAAAAACGGTAACTGCTATTAAAAAC 450
|||||
151  LLeuValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetas 167
|||||
451  ATTGTTATATTGGACAGCAAGGTGATTTAGAGGTTATCATCATGCA 500
|||||
167  PAspPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheI 184
|||||
501  CAACCTTTATTAATAAAACACTCCACAAAGTTTCAAAAGATCAAGTTTAA 550
|||||
184  ySThValAlaGluValAsnArgLysGluGlnValAlaLeuIleMetasSer 200
|||||
551  AAACGTGTGAAGTTAAACCGCAAGCAAGTTGCTTTATTAATGAACCTCT 600
|||||
201  SerGlySerThrGlyLeuProLysGlyValAlaGlnLeuThrHisGluAsnI 217
|||||
601  TCGGTTTACACCGGTTTCCAAAGGTGTGCACTTACTCATGAATAATT 650
|||||
217  eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
|||||
651  GGTACAGCGGTTTTCTCAGCTAGAGATCAATTTATGAAACCAAGTTT 700
|||||
234  eProGlyThrAlaIleLeuThrValIleProPheHisGlyPheGly 250
|||||
701  CACCGAGGCAAGGCTATTTTAACTGATGACATTCATCATGCTTTTGGT 750
|||||
251  MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
|||||
751  ATGTTTACTACTTATGCGTATCTAAGTGTGTTTCTGATTTGTCATGTT 800
|||||
267  uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
|||||
801  AACGAAATTTGACGAAAGACTTTTAAAAACACTGCAGATTTACAAT 850
|||||
284  ySerSerValIleLeuValIleProThrLeuPheAlaIleLeuAsnArgSer 300
|||||
851  GTTCAAGGCTTATTTCTTGACGACTTTGTTGCAATTTCTTAATGAGAGT 900
|||||
301  GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
|||||
901  GAATTAAGTGAATAATGATTTATCAAAATTTAGTTGAATTTGATCTGG 950
|||||
317  yGlyAlaProLeuSerLysGluIleGlyValAlaValAlaAspArgPheA 334
|||||
951  CGAGAGCACTTATCTAAAGAAATGTTGTAAGCTGTTGTAACGCTTTTA 1000
|||||
334  snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrSerAla 350
|||||
1001  ATTTACCGGGGTGTGCTCAAGGCTATGTTTAAAGAAACAACTCTGCA 1050
|||||
351  IleIleIleThrProGluLysAspLysProGlyAlaSerGlyLysVal 367
|||||
1051  ATTATTATATCACACCGGAGCGATGATTAACCAAGGCTTCGCAAAAGT 1100
|||||
367  ValProLeuPheLysAlaLysValIleAspLeuAspThrLysThrL 384
|||||
1101  TGTGCCATTATTTAAAGCAAAAGTTATCATCTCTGATCTATAAAAAAACTT 1150

```

```

384  eugLProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
|||||
1151  TGGGCCGAGACGACGCTGGAGAAAGTTTGTCTAAAGGCTCTTGTCTTAG 1200
|||||
401  LysGlyTyrValAspAsnProGluAlaThrArgGluIleLeuAspGluI 417
|||||
1201  AAGGTTATGATGATATATCCAGAACGACACAGAAAGAAATCATGATGAGA 1250
|||||
417  uGlyThrPheHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
|||||
1251  AGGTTGTTGACACAGAGATATTTGGTATTAAGTGAAGAAAAACATT 1300
|||||
434  hePheIleValAspArgLeuLysSerLeuIleLysTyrGlyTyrGln 450
|||||
1301  TCTTTATCGGATCGTTGAACTTTATATCAATACAAAGATATCA 1350
|||||
451  ValProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
|||||
1351  GTACCACTGCTGAATTAAGATCTGTTTGTGCAACATCCAAATATTTT 1400
|||||
467  eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
|||||
1401  TGATGCCGCGGCTGCTGGCTCCAGATCTTATAGCTGTGACCTTCGG 1450
|||||
484  LValAlaValAlaLeuLysLysGlyLysSerMetThrGluGluVal 500
|||||
1451  GAGCTGTTGTTGACTTGAAAAAGAAATCTATGACTGAAAAAGAAATTA 1500
|||||
501  MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgLysG 517
|||||
1501  ATGATTAAGCTGTGATGATCAAGTTTCAAAAGCAAAAGTTGCTGGTGG 1550
|||||
517  yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
|||||
1551  TGTCCGTTTGTGACGAACTACTTAAAGGCTCTCAGCTGTAATATTTGAGC 1600
|||||
534  LysAlaAlaIleArgGluIleLeuLysLysProValAlaLys 547
|||||
1601  GTAAAGCAATTAAGAAATTAAGTGAAGAAACCAAGTTGCTTAG 1641
|||||
seq_name: gb_pat:AR043316
seq_documentation_block:
LOCUS AR043316 1704 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 5 from patent US 5814465.
ACCESSION AR043316
VERSION AR043316.1 GI:5964324
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1704)
AUTHORS
Tatsumi, H., Fukuda, S., Kikuchi, M. and Koyama, Y.
TITLE
Biotinlated firefly luciferase, a gene for biotinlated firefly
luciferase, a recombinant DNA, a process for producing biotinlated
luciferase and a bioluminescent analysis method
JOURNAL
Patent: US 5814465-A 5 29-SEP-1999;
FEATURES
location/Qualifiers
1..1704
BASE COUNT 541 a 275 c 361 g 527 t
ORIGIN
alignment_scores:
Quality: 2799.00 Length: 546
Ratio: 5.126 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.267
alignment_block:
US-09-581-241-6 x AR043316 ..
Align seg 1/1 to: AR043316 from: 1 to: 1704

```

3 AsnMetCluAsnAspGluAsnIleValTyrGlyProGluProPheTyrPr 19
 : : : : :
 67 AGTCGAGAACGATGAAATATTGTGTATGTCCTGACCATTTTACC 116
 19 oileGluGluGlySerAlaGlyAlaGluLeuAglGlyTyrMetAspArgT 36
 : : : : :
 117 TATTGAGAGGATCTGCTGAGCACCAATGCGCAAGTATATGATGAT 166
 36 yzAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGlyValAsp 52
 : : : : :
 167 ATGCAGAACTTGAGCATGCTTTTACTACTACCATTTACCGGTGTCAT 216
 53 TyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGlyLysAlaLe 69
 : : : : :
 217 TATACCTACGCCCAATACCTAGAAAAATCATGCTGTAGAGAGGCTTT 266
 69 uysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCysSerGlu 86
 : : : : :
 267 AAGAATTATGCTTTGTTGTTGATGAGAAATTCGTTATGCGATGAAA 316
 86 sncCysGluLupPhePheIleProValLeuAlaGlyLeuPheIleGlyVal 102
 : : : : :
 317 ACTGTGAAAGTCTTATTTCTGTATGAGCGGTTATTTATAGGTGTC 366
 103 GlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuValHisSe 119
 : : : : :
 367 GGtGTGCTCAACTATAGATTACCTACCTACGGAATGTTGTTACAG 416
 119 rLeuGlyIleSerLysProThrIleValPheSerSerLysGlyLeuAl 136
 : : : : :
 417 TTTAGGCACTCTAAGCCAAACATTTATTTAGTTCTAAAAAGATTAG 466
 136 sPLysValIleThrValGluLysThrValThrAlaIleLysThrIleVal 152
 : : : : :
 467 ATAAAGTTATTAAGTACAAAAACCGTAACTGCTATTAACCATTTGTT 516
 153 IleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAspAsn 169
 : : : : :
 517 ATATTGGACGCAAGTGTATTAAGGTTATCAATTCATGACGACAACTT 566
 169 eileLysLysAsnThrProGlnGlyPheLysGlySerSerPheLysThrV 186
 : : : : :
 567 TATTAAAAAAACACTCCACAGAGTTTCAAGAGATCAAGTTTAAACAG 616
 186 aLGIuValAsnArgLysGluGlnValAlaLeuIleMetAsnSerSerGly 202
 : : : : :
 617 TAGAAGTTAAACGCAAGAACAGTTGCTTATATGAACTTCGCGGT 666
 203 SerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnIleValTh 219
 : : : : :
 667 TCAACCGGTTTCCAAAGGTGTGCACTTACTCATGAAAAATTTGGTCC 716
 219 rArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValSerProG 236
 : : : : :
 717 GCGTTTTCCTCAGCTAGAGATCCAAATTTATGAAACCAAGTTTACAGC 766
 236 lYThrAlaIleLeuThrValValProPheHisHisGlyPheGlyMetPhe 252
 : : : : :
 767 GCACGCGATTATTACTGTAGATCCATTCATCATGATGTTGGTATGTTT 816
 253 ThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLeuThrL 269
 : : : : :
 817 ACTACTTAAAGCTATCTAACTGTTGTTGTTGATTTGATGTTTAAACGA 866
 269 sPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysCysSerS 286
 : : : : :
 867 ATTGACAGAGAGACTTTTAAAAACATGCAAGATTACAAATGTCTCAA 916
 286 eValIleLeuValProThrLeuPheAlaIleLeuAsnArgSerGluLeu 302
 : : : : :
 917 GCGTTATTCTGTACCGACTTGTGCAATTTTAAATAGAAAGTGAATTA 966

303 LeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGlyGlyAl 319
 : : : : :
 967 CTCGATTAATATGATTTTATCAAAATTTAGTTGAATTCATCTGGCGAGC 1016
 319 aProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheAsnLeuP 336
 : : : : :
 1017 ACCTTATCTAAAGAAATGTTGTGAGAGCTGTTGCTAGACGTTTAAATT 1066
 336 roGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAlaIleIle 352
 : : : : :
 1067 CGGCTGTTCTCAAGGCTATGTTTAAACAGAAACACCTTGCATTAATT 1116
 353 IleThrProGluLysAspLysProGlyAlaSerGlyLysValAlaPr 369
 : : : : :
 1117 ATCACACCGGAGAGCGATGATTAACAGAGTCTTCTGCAAGTTGTGCC 1166
 369 oLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGlyP 386
 : : : : :
 1167 ATTATTTAAAGCAAAAGTATTCATCTTGCATCTAATAAAAAAATTGGGCC 1216
 386 roAsnArgArgGlyGlyValCysValLysGlyProMetLeuMetLysGly 402
 : : : : :
 1217 CGAACAGACCTGGAGAGATTTGTTAAAGGCTCTATGCTTATGAAGGT 1266
 403 TyrValAspAsnProGluAlaThrArgGluIleIleAspGluGluLyr 419
 : : : : :
 1267 TATGTAGATATATCCAGACAGCAACAGAAATCATGATGATGAAGAGT 1316
 419 PleuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisPhePhe 436
 : : : : :
 1317 GTTGACACAGAGAGATTTGGTATTACATGAGAAACAAATCTTTTAA 1366
 436 leValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGlnValPro 452
 : : : : :
 1367 TCGTGATGCTTTGAAGTCTTATATCAAAATFACAAAGATATCAAGTACA 1416
 453 ProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePheAspAl 469
 : : : : :
 1417 CCTGTGAATTTAGAAATCTGTTCTTTGCAACATCCAAATATTTTGATGC 1466
 469 aGlyValAlaGlyValProAspProIleAlaGlyGluLeuProGlyAlaV 486
 : : : : :
 1467 CGCGCTGCTGGCGTTCAGATCTATAGCTGTGAGCTTCGGGAGCTG 1516
 486 aValValLeuLysLysGlyLysSerMetThrGluLysGluValMetAsp 502
 : : : : :
 1517 TTGCTGTACTTGAAAGAGAAATCATGATGATGCAAAAAAACAATATGAT 1566
 503 TyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyValAla 519
 : : : : :
 1567 TAGGTTGCTAGTCAAGTTTCAATGCAAAACGTTGCGGTGGTGGTCCG 1616
 519 gPheValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLysA 536
 : : : : :
 1617 TTTTGTGAGAGAGATACCTAAAGGTCTCAGTGTAAATTTGACGGTTAAG 1666
 536 lAlleArgGluIleLeuLysLysProValAlaLysMet 548
 : : : : :
 1667 CAATTAGAGAAATACTGAAGAAACAGATTGCTAGATG 1704

seq_name: gb_pat:AR062708

seq_documentation_block:

LOCUS AR062708 1704 bp DNA PAT 29-SEP-1999

DEFINITION Sequence 5 from patent US 5843746.

ACCESSION AR062708

VERSION AR062708.1 GI:5990399

KEYWORDS Unknown.

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1704)

AUTHORS Tatsumi,H., Fukuda,S., Kikuchi,M. and Koyama,Y.

TITLE Biocinated firefly luciferase, a gene for biocinated firefly

luciferase, a recombinant DNA, a process for producing biocinated

luciferase and a bioluminescent analysis method

JOURNAL Patent: US 5843746-A 5 01-DEC-1998;

FEATURES Location/Qualifiers

1..1704

Source /organism="unknown"

BASE COUNT 541 a 275 c 361 g 527 t

ORIGIN

alignment_scores:

Quality: 2799.00 Length: 546

Ratio: 5.126 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 99.267

alignment_block:

US-09-581-241-6 x AR062708 ..

Align seg 1/1 to: AR062708 from: 1 to: 1704

```

3 AsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPheTyr 19
:.....
67 AGTCTCGAAGACGATGAAATATGTCTATGCTCGTCAACCATTTTACCC 116
19 oileGluGluGlySerAlaGlyAlaGluLeuAraGlyTyrMetAspArg 36
|||||
117 TATTGAAGAGGATCTGCTGAGCAGCAATGGCCAGATATATGATGAT 166
36 yralAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGlyValAsp 52
|||||
167 ATGCAAAACCTGGAGCAATTGCTTTACTAAGCAGCATTACCGGTCGAT 216
53 TyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGlyAlaIle 69
|||||
217 TATACGTAACGCCGCAATCTTAGAAAAATCAATGCTGAGAGAGGCTTT 266
69 uLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCysSerGlu 86
|||||
267 AAAGATATTATGTTGGTTGGTTGATGAGAGAAATTCGTTATGCGTAAA 316
86 snCysGluGluPhePheIleProValLeuAlaGlyLeuPheIleGlyVal 102
|||||
317 ACTGGAAGAGTCTTATTCCTGATTTAGCGGGTTATTATTATGAGGTC 366
103 GlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuValHis 119
|||||
367 GGTGGGCTCCAACTAATGAGATTACCTGACGTGATTTGTTCCACAG 416
119 rLeuGlyIleSerLysProThrIleValPheSerSerLysLysGlyLeu 136
|||||
417 TTTAGGCAATCTCTAAGCAACATTTGATTAGTTCTAAAAAAGAGATTAG 466
136 sPLysValIleThrValAlaGluLysThrValIleAlaIleLysThrIleVal 152
|||||
467 ATAAAGTTATACCTGTACAAAAACGGTAACTGCTATTTAAAAACATTGTT 516
153 IleLeuAspSerLysValAspTyrArgIlyTyrGlnSerMetAspAsp 169
|||||
517 AATTGGACAGCAAGTGAATTATAGAGTTATCAATCCATGAGCAACTT 566
169 eIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheLysThrV 186
|||||
567 TATTAAAAAAACACTCCACAAGTTTCAAGGATCAAGTTTAAACTG 616
186 aGluValAsnArgLysGluGluValAlaIleLeuIleMetAsnSerSerGly 202
|||||
617 TAGAGTTAAACGCAAGAAAGATTGCTTATATATGAAACTCTTCGGGT 666
203 SerThrGlyLeuProLysGlyValGlnLeuThrHisGlyAsnIleValThr 219
|||||
667 TCACCGGTTTGCAGAAAGGTGTGCACTTACTCATGAAAAATTTGGTCA 716

```

```

219 rArgPheSerHisAlaArgAspProIleTyrGlyAsnGluValSerProG 236
|||||
717 GCGTTTTCCTCAGCGTAGAGATCCATTTATGAAACCAAGTTTCACAG 766
236 lYThrAlaIleLeuThrValAlaProPheHisHisGlyPheGlyMetHe 252
|||||
767 GCACGGCTATTTTAACCTGATGATCATTCCATCAATGATGTTGGTATG 816
253 ThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLeuThr 269
|||||
817 ACTACTTAGGCTATCTAATCTTGTTGCTTTTCTGATTTGATGATTAACGA 866
269 sPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysCysSer 286
|||||
867 AATTGCAAGAGACTTTTAAAAACAGTCAAGATTAACAATGTCGAA 916
286 eValIleLeuValProThrLeuPheAlaIleLeuAsnArgSerGluLeu 302
|||||
917 GCGTTATCTTGTACCGACTTGTGTAATTTCTTAATGAAGTGAATTA 966
303 LeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGlyVal 319
|||||
967 CTCGATTAATATGATTTATCAATTTAGTTGAATTCGATCTGGCGAGAC 1016
319 aProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheAsnLeu 336
|||||
1017 ACCTTTATCTAAAGAAATGGTCAACGCTGCTGATGACCTTTTAAATTTAC 1066
336 rGlyValArgGlnGlyTyrGlyLeuThrGluThrSerAlaIleIle 352
|||||
1067 CCGGTTTGTCTCAAGCTATGTTTAAACAGAAACAACCTCTCAATTAAT 1116
353 lIeThrProGluGlyAspAspLysProGlyAlaSerGlyLysValAlaPr 369
|||||
1117 ATCAACACCGAAGCGATGATTAACAGAGCTTTCGGCAAGTTGTGCC 1166
369 oLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGlyP 386
|||||
1167 ATTTATTTAAAGCAAAAGTTATCGATCTGATTAATAAAAAACCTTGGGCC 1216
386 rAsnArgArgGlyGlyValAlaCysValLysGlyProMetLeuMetLysGly 402
|||||
1217 CGAACAGACCTGCAAGAGTTTGTAAAGGTCCTTATGCTTAAGAAAGT 1266
403 TyrValAspAsnProGluAlaThrArgGluIleIleAspGluGluLyr 419
|||||
1267 TATGTAGATTAATCCAGAACCAACAGAAATCATGATGAAGAAAGTTG 1316
419 pLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisPhePheI 436
|||||
1317 GTTGCACACAGAGATATTTGGTATTACGATGAAGAAACATTTCTTTA 1366
436 lValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGlnValPro 452
|||||
1367 TCCTGGATTCGTTGAGACTTTTAATCAAAATCAAGATTAACAGTACCA 1416
453 ProAlaGluLeuGluSerValIleLeuGlnHisProAsnIlePheAspAl 469
|||||
1417 CCGTCGAATTAAGATCTGTTCTTTTCAACATCCAAATATTTTGTATGC 1466
469 aGlyValAlaGlyValProAspProIleAlaGlyGluLeuProGlyValAl 486
|||||
1467 CCGCGTGTGCGCTTCAGATCCATATGCTGAGAGCTTCGCGGAGCTG 1516
486 aValValIleLeuLysLysGlyLysSerMetThrGluLysGlyValAlaMet 502
|||||
1517 TTGTTGTACTTGAAGAAAGAAATCTATGACTGAAAAAGAAATATGAT 1566
503 TyrValAlaSerGluValSerAsnAlaLysArgLeuArgGlyValAla 519
|||||
1567 TACGTGCTAGTCAAGTTTCAAAATGCAAAACGTTTGGCTGTGTCGCG 1616
519 gpHeValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLysA 536

```


353 IleThrProGluLysAspLysProGluAlaSerGlyLysValValPro 369
 1117 ATGACACCGGAGGCGATGATTAACAGGCTTGGCAAACTGTGCC 1166
 369 OleuPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGlyP 386
 1167 ATTATTTAAAGCAAAAGTTATCGATCTTGATTAATAAAAACTTTGGGCC 1216
 386 roAsnArgArgGlyGluValAlaLysValLysGlyProMetLeuMetLysGly 402
 1217 CGAACACACCTGGAGAGCTTGTGTAAAGGCTCTATGCTTATGAAGAGT 1266
 403 TyrValAspAsnProGluAlaThrArgGluIleIleAspGluGluGlyTr 419
 1267 TATGTAGATATATCCAGAACACAAAGAAATCATATGATGAGAGGCTTG 1316
 419 PLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisPhePheI 436
 1317 GTTCACACACGAGAGATTTGGGATTTACGATGAGAGAAAAACATTTCTTGA 1366
 436 leValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGlnValPro 452
 1367 TCGGAGATCGTTGAGCTTTATCAATACAAAGATATCAAGTACCA 1416
 453 ProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePheAspAl 469
 1417 CCGCTGTAATAGAACTGTTCTTTCCAAACATCAATATTTTGAATGC 1466
 469 aGlyValAlaGlyValProAspProIleAlaGlyLysLeuProGluValAla 486
 1467 CGGCTTCTGCTGCTCCAGATCCATATAGCTGAGAGCTTCCGGAGCTG 1516
 486 alValValLeuLysGlyLysSerMetThrGluLysGluValMetAsp 502
 1517 TTGTGTACTTGAAAAAGAAATCTATGACTGAAAAAGAGTATGAT 1566
 503 TyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyGlyValAla 519
 1567 TACGTGTGCTAGTCAGTTTCAATCAAAACCTTGGCTGGTGGTCCG 1616
 519 gPheValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLysA 536
 1617 TTTTGGGAGCAAGTAAAGCTCTACGTGTAATAATTCAGCGTAAAG 1666
 536 laIleArgGluIleLeuLysLysProValAlaLysMet 548
 1667 CAATTAGAGAAATCTGAGAAACCAAGTGTAGATG 1704
 seq_name: gb_pat:E13411
 seq_documentation_block: 1704 bp DNA PAT 24-JUN-1998
 LOCUS E13411 cDNA encoding biotinylated firefly luciferase bl203.
 DEFINITION E13411
 ACCESSION E13411
 VERSION E13411.1 GI:3252216
 KEYWORDS JP 1997187281-A/2.
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1704)
 AUTHORS Masuda,S., Tatsumi,H. and Koyama,T.
 TITLE ANTIBODY-FIREFLY LUCIFERASE FUSION PROTEIN. GENE OBTAINED
 THEREFROM. NOVEL RECOMBINANT DNA AND PRODUCTION OF ANTIBODY-FIREFLY
 LUCIFERASE FUSION PROTEIN
 JOURNAL Patent: JP 1997187281-A 2 22-JUL-1997;
 KIRKMAN CORP
 COMMENT OS Luciola lateralis (firefly)
 PN JP 1997187281-A/2
 PD 22-JUL-1997
 PF 09-JAN-1996 JP 1996001812
 PI MASUDA SUSUMU, TATSUMI HIROKI, KOYAMA TAJIJI
 PC C12N15/09,A61K39/395,C07H21/04,C07K14/47,C07K19/00,C12P21/02,

PC G01N33/535//
 PC C12N9/02,C12P21/02,C12R1:19);
 CC strandedness: Double;
 CC topology: Unknown;
 CC hypothetical: No;
 CC anti-sense: No;
 FH Key
 FT source Location/Qualifiers
 FT /clone='pHLf200' /organism='Luciola lateralis' FT
 FT mat_peptide 1..1704
 FT /product='Biotinylated firefly luciferase FT
 FEATURES
 source Location/Qualifiers
 1..1704
 /db_xref='taxon:32644'
 /organism='unidentified'
 /product='Biotinylated firefly luciferase FT
 BASE COUNT 541 a 275 c 361 g 527 t
 ORIGIN
 alignment_scores:
 Quality: 2799.00 Length: 546
 Ratio: 5.126 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.267
 alignment_block:
 US-09-581-241-6 x E13411 ..
 Align seg 1/1 to: E13411 from: 1 to: 1704
 3 AsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPheTyrP 19
 ::::|||||
 67 AGTCCTCAGACAGCATGAAATATTGTGTATGTCCTGACCAACATTTTACC 116
 19 oIleGluGluGlySerIleAlaGlnLeuArgLysTyrMetAspArg 36
 |||||
 117 TATTGAAGAGGAGTCTCTGAGCAGCATTTGCCAGATATATGATCGAT 166
 36 yAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGlyValAsp 52
 |||||
 167 ATGCAAACTTGGAGCATTTGCTTTACTTAACGACATTCGGGTGCGAT 216
 53 TyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGluAlaLe 69
 |||||
 217 TATACGTACGCCGGAATCTTAGAAAAATCATGCTGTCTAGAGAGGCTT 266
 69 ulAsnAspTyrGlyLeuValValAspGlyArgIleAlaLeuCysSerGlu 86
 |||||
 267 AAAGATTTATGCTTTGTTGTTGATGAGAGATTTGCTTATGACAGTGA 316
 86 snCysGluGluPhePheIleProValLeuAlaGlyLeuPheIleGlyVal 102
 |||||
 317 ACTGTGAAGAGTCTTTATTCCTGATTAAGCCGGTTTATTTATAGTGTC 366
 103 GlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuValHis 119
 |||||
 367 GGTGTGCTCCAACTAATGAGATTTACCTACGTACGTGATTTGTTCCAG 416
 119 rLeuGlyIleSerLysProThrIleValPheSerSerLysLysGlyLeu 136
 |||||
 417 TTACGACATCTCTAAGCCACAACTGATTAGTTCTTAAAAAGAGATTAG 466
 136 sPlysValIleThrValGlnLysThrValIleAlaIleLysThrIleVal 152
 |||||
 467 ATAAAGTTATTAATCTTACAAAAAAGGTAAGTCTATTAAAAACCACTTGT 516
 153 IleLeuAspSerLysValAspTyrArgLysTyrGlnSerMetAspAsn 169
 |||||
 517 ATATTGGACAGCAAGTGTGATTATAGAGTTATCAATCATGAGCAACTT 566
 169 eIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheLysThr 186

```

|||||
567 TATTAAAAAACACATCCACAGGTTTCAAGATCAAGTTTAAACAG 616
186 algluValasnArgLysGluGlnValAlaLeuIleMetLysSerGly 202
617 TAAAGATTAAACCCCAAGAACAGAGTGTCTATTAATGAACCTCTGGGT 666
203 SerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnIleValh 219
667 TCACACGGTTTGGCCAAAGGTGTGCAACTTACTCATGAAATTTGGTCCAC 716
219 ratGpSerHisAlaArgAspProIleTyrGlyAsnGlnValSerProG 236
717 GCGTTTCTCAGCTAGAGATCCAAATTTATGAAACCAAGTTTCAACAG 766
236 lyThrAlaIleLeuThrValAlaProPheHisHisGlyPheGlyMetPhe 252
767 GCACGGCTATTTTAACTGATAGTACCTTCATCATGCTTTTGGTATGTT 816
253 ThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLeuThrly 269
817 ACTACTTATGAGCTATCTTACTGTGTGTTTCCGTTATGTCATGTAAACGA 866
269 sPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysCysSerS 286
867 ATTGACGAGAGACTTTTAAAAACCTGCAGATTCACAAATGTCAA 916
286 erValIleLeuValProThrLeuPheAlaIleLeuAsnArgSerGluLeu 302
917 GCGTATTTCTGTACCGACTTGTTCGAATTCCTAATAGAAGTAAATTA 966
303 leuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGlyAla 319
967 CTGATTAATATGATTATTCAAATTTAGTTGAATTCATCTGGCGGAGC 1016
319 aProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheAsnLeuP 336
1017 ACCTTATCTAAAGAAATTTGGTGAAGCTGTGCTAGACGTTTAAATTAC 1066
336 roGlyValArgGlnGlyTyrGlyLeuThrGluThrSerAlaIleIle 352
1067 CGGGTGTCGTCACAGGCTATGTTTAAACGAACACCTCGCAATATAT 1116
353 IleThrProGluGlyAspAspLysProGlyAlaSerGlyLysValAlaPr 369
1117 ATCACACCGGAGGAGATGAATAAACCAAGCTCTCTGCAAGTTGTGCC 1166
369 oLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGlyP 386
1167 ATATATTTAAACCAAAAGTTATCGATCTTGATCTAATAAAAACTTGGGCC 1216
386 roAsnArgArgGlyGluValCysValLysGlyProMetLeuMetLysGly 402
1217 CGAACACAGCTGGAGAAAGTTTGTGTAAAGGCTCTATGCTTATGAAGGT 1266
403 TyrValAspAsnProGluAlaThrArgGluIleIleAspGluGlyTyr 419
1267 TATGTAGATATATCCAGAACCAAGAGAAATGATATAGTAAGAGTTG 1316
419 PLeuHisThrGlyAspIleGlyTyrTyrAspGluGlyLysHisPhePhe 436
1317 GTTGCCACACAGAGATTTGTGATTAACGATGAAGAAAAACATTTCTTTA 1366
436 leValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGlnValPro 452
1367 TCGTGGATCGTTTGAAGCTTTTATCAATACAAAGATATCAAGTACCA 1416
453 ProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePheAspAl 469
1417 CCGTCGAATTAGAATCTGTTCTTTGCAACATCCAAATATTTTTCATGC 1466
469 agLysValAlaGlyValProAspProIleAlaGlyGluLeuProGlyAlaV 486
|||||

```

```

1467 CCGCTTCTGGCGTTCCAGATCCTATAGCTGTGAGGTTCCGGAGCTG 1516
486 alValAlaLeuLysLysGlyLysSerMetThrGluLysGluValMetAsp 502
1517 TTGTGTACTTGAAGAAAGAAATCTATGACTGAAAAAGAAATGATGGAT 1566
503 TyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyValAla 519
1567 TACGTGCTAGTCAAGTTTCAATGCAAAACGTTTGGCTGGTGTCCG 1616
519 gPheValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLysA 536
1617 TTTGTGACCAAGTACTTAAGGTCTCAGCTGGTAAATTAATGAGTAAAG 1666
536 laIleArgGluIleLeuLysLysProValAlaLysMet 548
1667 CAATTAGAGAAATACTAGAGAACCACTGCTAGAGAG 1704

seq_name: em_pat:E10169
seq_documentation_block:
ID E10169 standard; DNA; UNC; 2019 BP.
XX
AC E10169;
XX
SV E10169.1
XX
DT 08-OCT-1997 (rel. 52, Created)
DT 02-SEP-2000 (rel. 65, last updated, Version 2)
XX
DE DNA encoding streptavidin-firefly luciferase fusion protein.
XX JP 1995289264 -A/1.
OS unidentified
OC unclassified.
XX
RN [1]
RP 1-2019
RA Tatsumi H., Fukuda M.;
RT "VARIANT TYPE STREPTAVIDIN GENE, FUSION PROTEIN GENE OF VARIANT TYPE
STREPTAVIDIN-FIREFLY LUCIFERASE, NEW RECOMBINANT AND PRODUCTION OF FUSION
PROTEIN OF VARIANT TYPE STREPTAVIDIN-FIREFLY LUCIFERASE",
RL Patent number JP1995289264 -A/1, 07-NOV-1995.
XX KIKKOMAN CORP.
XX
CC OS None
CC OC Artificial sequences.
CC PN JP 1995289264 -A/1
CC PD 07-NOV-1995
CC PE 27-APR-1994 JP 1994090275
CC PI TATSUMI HIROKI, FUKUDA MASARU
CC PC C12N15/09, C07K14/36, C07K19/00, C12P21/02, G01N33/53, C12P21/02,
CC C12R1:19);
CC CC strandedness: Double;
CC CC topology: Linear;
CC CC hypothetical: No;
CC CC anti-sense: No;
CC CC FH Key location/Qualifiers
CC CC FT source 1..2019
CC CC FT CDS 1..2019
CC CC FT /product="streptavidin-firefly luciferase
fusion protein"
XX
FH Key location/Qualifiers
FH FT source 1..2019
FH FT /db_xref="taxon:32644"
FH FT /organism="unidentified"
SQ Sequence 2019 BP; 590 A; 412 C; 467 G; 550 T; 0 other:

```

```

alignment_scores:      Quality: 2799.00      Length: 546
                       Ratio: 5.126      Gaps: 0
Percent Similarity: 100.000      Percent Identity: 99.267

alignment_block:
US-09-581-241-6 x E10169      ..

Align seg 1/1      to: E10169      from: 1      to: 2019

```



```

223 sAlaArgAspProIleuThyrGluAenGluValSerProGluThrAlaIleL 240
284 CTGAAGAGTGGGCTAAGCTGCTGGAGCTACCGGATCAGCTGGTACAC 333
240 eutThrValProPheHis.....GlyPheGlyMetPheThr 253
334 TACAGAGTACCTCTGTCTCCAGCTCCGCGGTGCGGCTG..... 377
254 ThrLeuGlyThrLeuPheGlyPheArgLeuValMetLeuThrLysPhe 270
378 ...CTGAATATGGTTACCAACAATGCCGTGTGTGTGATGCCACGATT 424
270 eaSPgluGluThrPheLeuGlyThrLeuGluAspTyrLysCysSerSery 287
425 TGATCCACACCTGTTCTTCTAGCTGATTAAGAATACGGCTAATCTAA 474
287 aAlaLeuValProThrLeuPheAlaIleLeuAsnArgSerGluLeuLeu 303
475 TGACACTGCTACCTCCGCTGATGCTCTTCTGGCCAAACATCCCATGCTG 524
304 AspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGlyGluAlaPr 320
525 GAAACTACATCTCTCTCTGCTGATGACCTGTTCTGTGGCGGCTCC 574
320 oleuSerLysGluIleGlyGluAlaValAlaArgArgPheAsnLeuProG 337
575 GTTGACCAAGAGATCGAAGATCAGTGGCGGAGCGACTGGTGTAGCTT 624
337 LysAlaArgGlnGlyThrGlyLeuThrGluThrPheSerAlaIleIle 353
625 TCATTGGCAAGGTGACGATGATGAAACCAACGCTCGGTGTGTGATG 674
354 ThrProGluGlyAspAspLysProGluAlaSerGlyLysValAlaPro 370
675 CAGAGCGATTCGAAACAAGCGGCGACGCTGGGAAGTGGCGCATGGG 724
370 uPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGlyPro 387
725 CCAGTGGGGAAGGTGATGACCCGCGAGCGGCGCTGCGCTGCAGCGA 774
387 snArgArgGlyGluValLysValLysGlyProMetLeuMetLysGlyTyr 403
775 ACCAGCGGGGTGAGCTGCTTTAAGGCTCATTTGATGAAGAATAT 824
404 ValAspAsnProGluAlaThrArgGluIleLeuAspGluGlyTyrPhe 420
825 GTGGCGCAAGACGAGCC.....ATCGATGCTGACGCTGGCT 862
420 uHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisPhePheIle 437
863 GCACAGGGGCGACATTCCTACTACATGATGATCAGAGTTTTCATCG 912
437 aAlaAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGlnValPro 453
913 TCGACCGGATCAAGAGCTGATCAAGTCAAGGCTTCCARGTACCGCG 962
454 AlaGluLeuGlu 457
963 GCCGAGCTGGAA 974

seq_name: gb_est2:BG445463
seq_documentation_block: 882 bp mRNA EST 15-MAR-2001
LOCUS BG445463
DEFINITION GA_Ea0028K14f Gossypium arboreum 7-10 dpa fiber library Gossypium
arboresum cDNA clone GA_Ea0028K14f, mRNA sequence.
ACCESSION BG445463
VERSION BG445463.1 GI:13355211
KEYWORDS EST.
SOURCE Gossypium arboreum.
ORGANISM Gossypium arboreum.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

```

Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots;
Rosidae: eurosids II; Malvales; Malvaceae; Gossypium.
1 (bases 1 to 882)
Wing, R.A., Fritsch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry
D., Wood, T.C., Leslie, A. and Wilkins, T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished (2000)
CONTACT: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATACACTCATATAGCG
High quality sequence stop: 733.
Location/Qualifiers
1..882
/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0028K14f"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/Note="vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 250 a 173 c 232 g 226 t 1 others
ORIGIN

alignment_scores:
Quality: 552.00 Length: 277
Ratio: 2.816 Gaps: 4
Percent Similarity: 70.758 Percent Identity: 41.155

alignment_block:
US-09-581-241-4 x BG445463 ..

Align seg 1/1 to: BG445463 from: 1 to: 882
263 ArgIleValMetLeuThrLysPheAspGluGluThrPheLeuLysThrLe 279
|||||
6 CGAGTTTATCATGACGAAAGTTCGACATCGGCTACTGTGAGCTAAT 55
|||||
279 uGlnAspTyrLysCysSerSeryAlaIleLeuValProThrLeuPheAla 296
|||||
56 ACAGAAATACAAAGTAAACCATTCCTCCGATTCGCCACCGATGTTTGG 105
|||||
296 LeuLeuAsnArgSerGluLeuLeuAspLysTyrAspLeuSerAsnLeuVal 312
|||||
106 CCATTCCTAAGTCAATCGAAACTGAAATTAACGATTTGTCATCGGAGG 155
|||||
156 ATGTGAAGTCCGCTCCGCTCGTTGGGTCAAGAGCTTGAGAGATGCTCT 205
|||||
329 LAlaArgArgPheAsnLeuProGlyValArg.....GlnGlyTyrGlyL 344
|||||
206 AAAGTCAAGTTT.....CTGTGTCCAAAATTGGACAGGGGTTATGGA 249
|||||
344 eutThrGluThrSerAlaIleIleIle.....ThrPro 355
|||||
250 TCACAGAACTGACACAGTTCACAAATGTGTTGGATTTCCCAAGCA 299
|||||
356 GluGlyAspAspLysProGlyAlaSerGlyLysValAlaProLeuPheLys 372
|||||
300 CCTTTGAATGAATTCGGAACCTTGTGGAGCTGTTGAAGAACGCGGA 349
|||||
372 sAlaLysValIleAspLeuAspThrLysLysThrLeuGlyProAsnArg 389
|||||

```

seq_name: gb_est1:BE036528	seq_documentation_block:	LOCUS	BE036528	1066 bp	mRNA	EST	07-JUN-2000
DEFINITION	MP01AN2 MP Mesembryanthemum crystallinum cDNA 5' similar to	ACCESSION	BE036528	VERSION	KEYWORDS	comarate:coenzyme a ligase, mRNA sequence.	
SOURCE	ORGANISM	common ice plant. Mesembryanthemum crystallinum Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Caryophyllales; Caryophyllales; Aizoaceae; Mesembryanthemum. 1 (bases 1 to 1066) Bohnerl,H.J., Borchert,C., Brazille,S., Brooks,J., Eaton,M., Ferreira H., Kawasaki,S., McCollough,A., Michalowski,C.B., Palacio,C., Scera,G., Wheeler,M. and Zepeda,G.R.					
TITLE	JOURNAL	COMMENT	Functional Genomics of Plant Stress Tolerance Unpublished (2000) Contact: Michalowski,C.B. University Of Arizona Bio Sciences West room 513, Tucson, AZ 85721, USA Tel: 520-621-7982 Fax: 520-621-1697 Email: cmheu.arizona.edu An open reading frame exists. Insert Length: 1 Std Error: 0.00. Location/Qualifiers 1..1066 /organism="Mesembryanthemum crystallinum" /db_xref="taxon:3544" /clone_lib="MP" /tissue_type="apical meristem and leaf primordia" /dev_stage="6 weeks"				

BASE COUNT	265 a	220 c	279 g	294 t	8 others			
ORIGIN	/note="3 d 500mM NaCl"							
Alignment_scores:	Quality: 541.00	Length: 295						
	Ratio: 2.678	Gaps: 5						
Percent Similarity:	68.475	Percent Identity: 40.339						
Alignment_block:								
US-09-581-241-4 x BE036528	..							
Align seg 1/1 to: BE036528 from: 1 to: 1066								
218	ValThrArgPhe.....SerHisAlaArgAspProIleThrGlyAsnG1	232	295				
4	ATTACGGCGCTATAGGAGTCACTATTAATCTCCG.....	39	5				
232	nValSerProGlyThrAlaIleLeuThrValValProPheHisGlyP	249	5				
40CCAAAGGGCGCTGTTCTTCCTCCGTTTCATGCTCT	76	5				
249	heGlyMetPheThrThrLeuGlyThrLeuThrGlyPheArgIleVal	265	5				
77	TGGTTCTTCCTGAGTCTCTGAGAGGGGTGTTGGCCAAACTAGTT	126	5				
266	MetLeuThrLysPheAspG1uG1uThrPheLeuLysThrLeuGlnAsp	282	5				
127	GTAATGGCAAGGTGATTCGAGAAAGATGTTGAGGCTGGAGGAAGTA	176	5				
282	rlYscYsSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnA	299	5				
177	TCGGATTACATCAATTCGGGTTCCGCCCATTTGGTGTGACATGCAAA	226	5				
299	rgSerGluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAla	315	5				
227	AATCGAGCTTACATAGTAAGTACATCTCACACAGCTGGAAAGTCTGCT	276	5				
316	SerGlyGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgA	332	5				
277	TGAGTGGGCGCTCCTCTGGAGAGGGTTGCACAGGGTTTCGAGCTCG	326	5				
332	grPheAsnLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrAs	349	5				
327	CTTCCCGAAGCTGGAGATAGCACAGGAGTATGGCTGACTAGACAGCAG	376	5				
349	erAlaIleIleIleThr.....ProGluGlyAspAspLysProGlyAla	363	5				
377	GGCGGCGCTCAAAAGACGTTAGGGCTCTGAGCATGTGAACGATATGGTCT	426	5				
364	SerGlyLysValValProLeuPheLysAlaLysValIleLeuLeuAspTh	380	5				
427	GTTGGCGGCTTACAGAGTCTTGGAGACCAAGATAGTGAATCCGCTAAC	476	5				
380	rlYsLysThrLeuGlyProAsnArgArgGlyGluValAlcYsValLysGlyP	397	5				
477	CGAGAGAGGCTTACACACCTGTCGCAAGAGGGGAGCTGTGGCTGGGGC	526	5				
397	roMetLeuMetLysGlyTyrValAspAsnProGluAlaThrArgGluIle	413	5				
527	CAACCAATATGAANAACCTATGTGGAGATAAGAGATGCCACGTGCAACT	576	5				
414	IleAspG1uG1uTyrPheLysThrGlyAspL1leGlyTyrTyrAspG1	433	5				
577	NTCCATTCANAAAGATGGTTGAAGACTGTGTATCTTTGATTTTGACTC	628	5				
430	uGluLysHisPhePheIleValAspArgLeuLysSerLeuIleLysTyrL	447	5				
627	TGATGGCTTTCTTCATATGTTGATCGATTAAAGCAATNGATTAAGTACA	676	5				
447	ysGlyTyrGluValProPheAlaGluLeuGluSerValLeuLeuGlnHis	463	5				

```

677 AGGATTCACNGTCCCTCCAGCTGAGTTGGAGCGGTTCATCATCATTCCAT 726
464 ProasnlIerheaspAlagIyValAlaGIyValProaspProIIealGl 480
|||||:||||| ||||| ||||| |||||
727 CCAGATATTCGTGTAGTGCTGCCTGTGTGCCGTATCCAGATACCAAGCNG 776
|||||:||||| ||| ||||| ||||| |||||
480 yGUluePrGclYAlaValAlaVallleuLysGLy.LysSerMetThr 496
|||||:||||| ||| ||||| ||||| |||||
777 ACAATATNCCTATGSCATTGTGTGTC..AGGAATNCAAGAACCACTTTC 823
|||||:||||| ||||| ||||| |||||
seq_name: gb_est1:AU214129

seq documentation_block:
LOCUS AU214129 741 bp mRNA EST 17-JUL-2001
DEFINITION AU214129 unpublished oligo-capped cDNA library, stage L2
ACCESSION Caenorhabditis elegans CDNA clone YK809f05 3', mRNA sequence.
VERSION AU214129
KEYWORDS AU214129.1 GI:14852286
SOURCE EST
ORGANISM Caenorhabditis elegans.
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea
1 (bases 1 to 741)
Koehara,Y., Shln-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
A complementary view of the C.elegans genome
Unpublished (2001)
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
FEATURES
source
1..741
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk809f05"
/cds_1lb="unpublished oligo-capped cdna library, stage
L2"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L2"
BASE COUNT 216 a 158 c 125 g 241 t 1 others
ORIGIN

alignment_scores:
Quality: 513.00 Length: 223
Ratio: 3.090 Gaps: 5
Percent Similarity: 74.439 Percent Identity: 47.982

alignment_block:
US-09-581-241-4 x AU214129/rev ..

Align seg 1/1 to reverse of: AU214129 from: 1 to: 741

320 ProteuserysGUllieglYUlaValAlaArgaPhe..AsnLe 335
|||||:|||||:||||| ||| :|||:||||| |||||
739 CCAGCTGGAAAGGATTTGATTGAGAATTTGAAGAGAGATATACAAATTT 690
|||||:|||||:|||||:|||||:|||||:|||||
335 uProgiyValArgGInglyTyrglyLeuThrgIGurhTrSerAlaIei 352
|||||:|||||:|||||:|||||:|||||:|||||
689 GAATATACATCAACAAGATATGATGATGCGAATGTCATGCGCTAGTC 640
|||||:|||||:|||||:|||||:|||||:|||||
352 IelleThProGUclYspasplysPro..gIyAlaserGlylysVal 367

```

seq_name: qb_estL2:C83857

seq_documentation_block:

LOCUS C83857 664 bp mRNA EST 28-APR-1999

DEFINITION C83857 Dictyostelium discoideum SS (H. Urushihara) Dictyostelium

DISCOIDEUM cDNA clone SSA391, mRNA sequence.

ACCESSION C83857

VERSION C83857.1 GI:2706789

KEYWORDS EST.

SOURCE Dictyostelium discoideum.

ORGANISM Dictyostelium discoideum.

Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.

1 (bases 1 to 664)

Morita,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M., Yoshino,R., Mita,B.N., Pl.M., Sato,T., Takemoto,K., Yasukawa,H., Williams,J., Maeda,M., Takeuchi,I., Ochiai,I. and Tanaka,Y.

The Dictyostelium developmental cDNA project: generation and analysis of expressed sequence tags from the first-finger stage of development

DNA Res. 5 (6), 335-340 (1998)

99156227

CONTACT: Hideko Urushihara

Institute of Biological Sciences

University of Tsukuba

3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan

Email: dd02h@tsukuba.ac.jp

PROJECT = 'Dictyostelium discoideum cDNA project in Japan'
POLYA-No.

FEATURES
source location/Qualifiers
1.664
/organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44689"
/clone_lib="SSA391"
/dev_stage="slug"
/clone_stage="slug"

BASE COUNT 234 a 105 c 124 g 201 t
ORIGIN

alignment_scores:
Quality: 512.00 Length: 200
Ratio: 3.261 Gaps: 0
Percent Similarity: 78.500 Percent Identity: 47.500

alignment_block:
US-09-581-241-4 x C83857 ..

Align seg 1/1 to: C83857 from: 1 to: 664

```

340 GlnGlyTyrGlyLeuThrGluThrSerAlaIleIleIleThrProG1 356
|||||
2 CAAGGTATGGTCTACCGAATAGTCCAGCGCTTTGTAATTCAG 51
|||||
356 uGlyAspAspLysProGlyAlaSerGlyLysValAlaProLeuPheLysA 373
|||||
52 TGGATTAGTAAATCTGGTCCGCTGATCCCTTACCACATCAATGG 101
|||||
373 lAlaValIleLeuPheLysPheLysThrLeuGlyProAsnArgArg 389
|||||
102 CTAAATATCTCTCCAGAGACTGGTGGAATTTAGTATGGGTAAG 151
|||||
390 GlyGluValCysValLysGlyPrometLeuMetLysGlyTyrValAspAs 406
|||||
152 GGTGAATCTGTATTAAGTCCAAATGTTATGTTGTTATTAACA 201
|||||
406 nProGluAlaThrArgGluIleIleAspGluGluGlyTyrPheLysThrG 423
|||||
202 TCAGAAAGGCAACCAATGAATCATGAATTAAGATGATTTAAAAACTG 251
|||||
423 lAspIleGlyTyrTyrAspGluGluLysHisPhePheIleValAspArg 439
|||||
252 GTGATATTGGTATCGTCGATGAAGATGGTATTACTTATCGTCATAGA 301
|||||
440 LeuLysSerLeuIleLysTyrLysGlyTyrGluValProProAlaGluLe 456
|||||
302 TCAGAAAGGCAATGAATGAATGTTCCAAAGTACCTCCGCGCAAT 351
|||||
456 uGlySerValIleLeuGluHisProAsnIlePheAspAlaGlyValAlaG 473
|||||
352 GGAGGCAATTCATTAATCATCAAAAGTTCAGATGCTGTGTAGTAG 401
|||||
473 lValProAspProIleAlaGlyGluLeuProGlyAlaValAlaValLeu 489
|||||
402 GTCTTCAAAAGGTATAGGTGTAAGTCAAGTCAAGAGGTTCTCTGTAT 451
|||||
490 LysLysGlyLysSerMetThrGluLysGluValMetAspTyrValAlaSe 506
|||||
452 AAAAATAATGAATCTCTAACGAAAAAAGCAACTTATGATGGGCCATCC 501
|||||
506 rGluValSerAsnAlaLysArgLeuArgGlyGlyValAlaArgPheValAsp 523
|||||
502 AAAAATTCGAAATTAATTAACATTTTCAGAGGTGTATCTTTTCATACCTG 551
|||||
523 lValProLysGlyLeuThrGlyLysIleAspGlyLysAlaIleLeuArgL 539
|||||
552 CATTTCTTAATCTGCACCGGTAACATTAATTCGTAATAAATCTTAAGAT 601
|||||

```

seq_name: gb_est2:BF492687

seq_documentation_block:

LOCUS BF492687 683 bp mRNA EST 19-APR-2001
DEFINITION AT01136.3prime AT Drosophila melanogaster adult testes potB7.
Drosophila melanogaster cDNA clone AT01136.3 similar to CG6178.
Fban0006178 enzyme located on: 3R 95E1-95E1.1: 04/07/2001, mRNA
sequence.

ACCESSION BF492687
VERSION BF492687
KEYWORDS EST

SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 683)
Stapleton, M., Brooks, P., Hong, L., Aghayani, A., Baxter, E., Berman, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V., Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Miranda, A., Misra, S., Mungall, C.J., Nuno, J., Pacle, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Celniker, S. and Rubin, G.M.
BDGP/HM1 AT Drosophila EST Project
Unpublished (2000)
Other ESTs: AT01136.5prime
Contact: Stapleton, M.
BDGP

TITLE BDGP/HM1 AT Drosophila EST Project
JOURNAL Unpublished (2000)
COMMENT Other ESTs: AT01136.5prime
Contact: Stapleton, M.
BDGP

One Cyclotron Rd
Lawrence Berkeley National Lab
Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_estfruitfly.berkeley.edu
hit genomic AB003746: arm:3R [19764572..19982668]
estimated-cyto:95D5-95F1: 04/07/2001
Plate: AT.11 row: C column: 12
High quality sequence strop: 631.

FEATURES

source

Location/Qualifiers

1..683
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="AT01136"
/clone_lib="AT Drosophila melanogaster adult testes potB7"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DHS-alpha. Plates AT.121-AT.319: DHS-alpha Tona"
/note="Organ: ADULT testes; Vector: potB7; Site:1: EcORI; Site:2: XhoI; The mRNA for the testis library was made from testes and seminal vesicles hand dissected from 0-3 day old Ore-R males. RNA kindly provided by the lab of Margaret Fuller. Sized fractionated cDNAs were directly ligated into potB7. Plasmid cDNA library."

BASE COUNT 169 a 188 c 149 g 177 t
ORIGIN

alignment_scores:
Quality: 507.50 Length: 176
Ratio: 3.574 Gaps: 1
Percent Similarity: 80.682 Percent Identity: 57.955

alignment_block:
US-09-581-241-4 x BF492687/rev ..

Align seg 1/1 to reverse of: BF492687 from: 1 to: 683

```

373 AlAlaValIleAspLeuAspThrLysLysThrLeuGlyProAsnArgArg 389
|||||
637 GCGACAGGTGATGATCCGACACCGGCAAGCTATTGGGGCCCAACGAGCG 588
|||||
389 gGlyGluValCysValLysGlyPrometLeuMetLysGlyTyrValAspA 406
|||||
587 CGGCGAGCTTTGTTTAAAGCGACGCGATCATGAAGGCGTACATCGGAG 538
|||||

```

```

406 snpGluAlaThrArgGluIleLeuAspGluGluGlyTyrPheHisThr 422
      ::::::::::::::|::|::::::::::::|
537 ATGAGAGTCCAGCAGACCCGATC...AAGAGCGGTGGTTCATACT 491
      ::::::::::::::|::|::::::::::::|
423 GLAAspIleGlyTyrAspGluGluHisPheIleValAspArg 439
      |||
490 GGGGATTTGGCTACTATGATGATGATTTGAGTTCTTCATCGTGACCG 441
      |||
439 GLeuYSerLeuIleLysTyrGlyTyrGlnValProPheIleVal 456
      ::::::::::::::|::|::::::::::::|
440 CATTCAAGAGCTGATCAATACAAAGGATACAGAGTCCCGCCGACAGA 391
      ::::::::::::::|::|::::::::::::|
456 euGluSerValLeuLeuGlnHisProAsnIlePheAspAlaGlyValAla 472
      ::::::::::::::|::|::::::::::::|
390 TTGAGGCTCGCTGCTCCACCAAGATTTAAGATGGCGCGCATAT 341
      ::::::::::::::|::|::::::::::::|
473 G1ValProAspProIleAlaGlyGluLeuProGlyAlaValValLe 489
      |||
340 GGAAGGCCAAGCAGAGAGGCTGGCGAGCTGCCGCTGATTTGTCTAA 291
      ::::::::::::::|::|::::::::::::|
489 uLysLysGlyLysSerMetThrGluGluValMetAspTyrValAla 506
      ::::::::::::::|::|::::::::::::|
290 ACAGGCTAATGTCATCACTGACCGAGAACGAAATCATCTGTTGTCAAC 241
      ::::::::::::::|::|::::::::::::|
506 erGlnValSerAsnAlaLysArgLeuArgGlyValArgPheValAsp 522
      ::::::::::::::|::|::::::::::::|
240 ACACGCGCTCGCCGACGAGCTTAAGGGGCTGATCTTGTGTGAC 191
      ::::::::::::::|::|::::::::::::|
523 G1uValProLysGlyLeuThrGlyLysIleAspGlyLysAlaIleArg 539
      |||
190 GAATTCGCAAGAAAGAACCCAGTGGCAAGATTCGCGTCGATTCGCGGA 141
      ::::::::::::::|::|::::::::::::|
539 uIleLeuLysLysProValAlaLysMet 548
      ::::::::::::::|::|::::::::::::|
140 AATGCTTAAGAACCAAAATCAAGTTG 113

seq_name: gb_est1:A1486799

seq_documentation_block:
LOCUS      A1486799      667 bp      mRNA      EST      18-MAY-2001
DEFINITION EST245121 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
            CLEDD1D3, mRNA sequence.
ACCESSION  A1486799
VERSION    A1486799.1 GI:4382170
KEYWORDS   EST.
SOURCE     tomato.
            Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; eusterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
            1 (bases 1 to 667)
REFERENCE  1 (bases 1 to 667)
AUTHORS   Alcalá, J., Vrebalov, J., White, R., Matern, A. L., Vision, T., Holt, I. E.,
            Liang, F., Upton, J., Romling, C. M., Craven, M. B., Fujii, C. Y., Bowman,
            C. L., Mierman, W., Fraser, C. M., Venter, J. C., Martin, G. B., Tanksley,
            S. D. and Giovannoni, J.
            Generation of ESTs from tomato carpel tissue
            Unpublished (1999)
COMMENT    CUGI
            Clemson University Genomics Institute
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html.

FEATURES
            source
            1..667
            /organism="Lycopersicon esculentum"
            /cultivar="T496"
            /db_xref="taxon:4081"
            /clone="CLEDD1D3"
            /clone_lib="tomato ovary, TAMU"
            /tissue_type="carpel"
            /dev_stage="5 days pre-anthesis to 5 days post-anthesis"
            /lab_host="X11-Blue MRP"

```

```

/Note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; CLEDD - Tomato Carpel EST library. OligodT-primed and
directionally cloned cDNA in vector Lambda Zap II with 5'
and 3' ends located at the EcoRI and XhoI sites,
respectively."
BASE COUNT      201 a      120 c      155 g      191 t
ORIGIN
alignment_scores:
      quality: 505.50      length: 222
      Ratio: 3.159
      Percent Similarity: 72.072      Percent Identity: 46.396
alignment_block:
US-09-581-241-4 x A1486799 ..
Align seq 1/1 to: A1486799 from: 1 to: 667
280 GlnAspTyrLysCysSerSerValIleLeuValProThrLeuPheAlaI 296
      |||
1 CAAATTCAGAGGAGGACCATAGGCGCATTTGTGCGACCTATGTTTGGC 50
      ::::::::::::::|::|::::::::::::|
296 eLeuAsnArgSerGluLeuLeuAspLysTyrAspLeuSerAsnLeuVal 313
      ::::::::::::::|::|::::::::::::|
51 TATTGCTAAGAGTCCTATGCTGATATTAATTAATTAATTCATCGGTAGA 100
      ::::::::::::::|::|::::::::::::|
313 IuIleAlaSerGlyGlyAlaProLeuSerGlyIleGlyAlaVal 329
      ::::::::::::::|::|::::::::::::|
101 CCGTTATGTCTGGGCTGCACCATTAATTAAGAAAGAACTTGAAGACAGT 150
      ::::::::::::::|::|::::::::::::|
330 AlaArgPheAsnLeuProGlyValArgGlnGlyTyrGlyLeuThr 346
      ::::::::::::::|::|::::::::::::|
151 CGAGCCAAATTTCTTAATGCTAAACTTGTCAGAGTTACGGAATGACGA 200
      ::::::::::::::|::|::::::::::::|
346 uThrThrSerAlaIleIleIle.....ThProGluGlyA 358
      ::::::::::::::|::|::::::::::::|
201 AGCCGACAGAGTGTGCTATGCTTGGCATTTGCTAAAGAACCCCTTCG 250
      ::::::::::::::|::|::::::::::::|
358 sPaSpLysProGlyAlaSerGlyLysValAlaProLeuPheLysAlaLys 374
      ::::::::::::::|::|::::::::::::|
251 AATTTAAATTTGGAGCAATGGGACCTGTTGTAAGAAATGCTCAGATGAAA 300
      ::::::::::::::|::|::::::::::::|
375 ValIleAspLeuAspThrLysLysThrLeuGlyProAsnArgArgGly 391
      ::::::::::::::|::|::::::::::::|
301 ATTGTGATTCCTGATCTAGTGGCAATCTCTCCCTAGAACCAATCTGGAGA 350
      ::::::::::::::|::|::::::::::::|
391 uValLysValLysGlyProMetLeuMetLysGlyTyrValAspAsnPro 408
      ::::::::::::::|::|::::::::::::|
351 GATTTGTATTAAGAGCGGCAATCAATTTATGAAGGTTACTCAATGATCCAG 400
      ::::::::::::::|::|::::::::::::|
408 IuAlaThrArgGluIleIleAspGluGluGlyTyrPheHisThrGlyAsp 424
      |||
401 AGGCCACTACGGGACAAATAGCAAGAAAGAGGTGTACTACAGCGCGAT 450
      ::::::::::::::|::|::::::::::::|
425 IieGlyTyrTyrAspGluGluLysHisPheIleValAspArgLeu 441
      |||
451 ATTGTTTATTAATTCATCATGATGATGATTTGATGATGATGATTTAA 500
      ::::::::::::::|::|::::::::::::|
441 sSerLeuIleLysTyrLysGlyTyrGlnValProPheAlaGluLeuGln 458
      ::::::::::::::|::|::::::::::::|
501 GGAATTCATTAATTAACAAAGATTTCAAGTGGCTCGCTGCAACTGAAG 550
      ::::::::::::::|::|::::::::::::|
458 erValLeuLeuGlnHisProAsnIlePheAspAlaGlyValAlaGlyVal 474
      ::::::::::::::|::|::::::::::::|
551 CCTTCTCCTCAATCATCCCAATATTTCAGATGCTGCTTGTTCATAG 600
      ::::::::::::::|::|::::::::::::|
475 ProAspProIleAlaGlyLeuProGlyAlaValValLeuLysLys 491
      |||
601 AAAGACGACCAAGCAGAGAGTCTCTGCGCTTGTGTGATCAGATCAAA 650
      ::::::::::::::|::|::::::::::::|
491 sGlyLysSerMetThr 496
      ::::::::::::::|::|::::::::::::|

```

651 TGGCTCCACAATTACT 666

seq_name: gb_est2:BG414540

seq_documentation_block:

LOCUS BG414540 876 bp mRNA EST 13-MAR-2001

DEFINITION HVSMEK0002N03f Hordeum vulgare testa/pericarp EST library

HYCDNA0013 (normal) Hordeum vulgare cDNA clone HVSMEK0002N03f, mRNA sequence.

ACCESSION BG414540

VERSION BG414540.1 GI:13320091

KEYWORDS EST.

SOURCE barley.

ORGANISM Hordeum vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Hordeum.

REFERENCE 1 (bases 1 to 876)

Wing, R., Close, T.J., Kleinjohs, A., Wise, R., Begum, D., Frisch, D., Yu, Y., Anderson, H., Dale, J., Henry, D., Kernodle, S., Palmer, M., Rambo, T., Saski, C., Schwartzbeck, J., Simmons, J., Choi, D.W., Main, D. and Wood, T.

Development of a genetically and physically anchored EST resource for barley genomics

UNPUBLISHED (2000)

CONTACT: Wing RA

Clemson University Genomics Institute

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: rwing@clemson.edu

Seq primer: AATTAACCTCCTACTAAGG

High quality sequence stop: 847.

Location/Qualifiers

1..876

/organism="Hordeum vulgare"

/cultivar="Morex"

/db_xref="taxon:4513"

/clone="HVSMEK0002N03f"

/clone_id="Hordeum vulgare testa/pericarp EST library

HYCDNA0013 (normal)"

/issue_type="testa/pericarp"

/lab_host="TJC121"

/note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI; For more details on library preparation and sequence analysis see <http://www.genome.clemson.edu/projects/barley/> to order a clone see <http://www.genome.clemson.edu/orders>"

BASE COUNT 248 a 187 c 209 g 232 t

ORIGIN

alignment_scores:

Quality: 505.50 Length: 215

Ratio: 3.082 Gaps: 4

Percent Similarity: 76.279 Percent Identity: 49.767

alignment_block:

US-09-581-241-4 x BG414540 ..

Align seg 1/1 to: BG414540 from: 1 to: 876

323 LysGluIleGlyGluIleValAlaArgArgPheAsnLeuProGlyValAr 339

6 AGGAGGCTCATGAGGTGGTGGCCAGAACTTCCCGACGCCGAGATCGT 55

339 gGlnGlyTyrGlyLeuThrGluThrThrSerAlaIleIleIleThrProg 356

56 TCAGGGCTATGATAGCTGAGACTGTGGG.....ATCATATCGCTTG 99

356 LucGlyAspAspLysPro.....GlyAlaSerGlyLysVal 367

100 AATTAACCAAGAAAGGACAGGCTGCTCAATTGGGTCAACCGAAGCTT 149

368 ValProLeuPheLysAlaLysValIleAspLeuAspThrLysThrIle 384

150 GTTGTGGAGTTGAGCAAAAAGTATATGACGTAGAAAACGCAAAACATCT 199

384 uGlyProAsnArgArgGlyGluValLysValLysGlyPrometLeuMetL 401

200 ACCGCCAAGTCACTAGAGAAATCTATCCGAGCACACATATATGC 249

401 ySgIlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGluIu 417

250 AAGGCTATTTCACAAATGATAGGCTACT...GAATTACAAATCGCAA 296

418 GLyTrpLeuHisThrGlyAspIleGlyTyrTyrAspGluLysHisph 434

297 GGATGGTTGCTACTGGTGTGATCTTGATCTTGATTAAGAACGCCAGCT 346

434 ePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGlnv 451

347 TTTGTGCTGCTAGTAAAGAGCTGATTAAAGTACAAAGTTCCAGA 396

451 aLProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePhe 467

397 TTGCACCTCGCTGAGCTTGAGAGATTCTTATCTCATTCACAGATTCTA 446

468 AspAlaGlyValAlaGlyValAlProAspProIleAlaGlyLysLeuProG1 484

447 GAGCCGCGCTTATCCGCTTCCGTGATGATGAAGCTGGGAGTTCCTAT 496

484 yAlaValAlaValLeuLysLysGlyLysSerMetThrGluLysGluValM 501

497 TGCTATGTGGTGAAGTCACTGCCAGTCAATTAACGGAAGTGCAGCTCC 546

501 eLAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyGly 517

547 AAAATTATGCGCAACAGGTGACGACTACAAACAAAGCTGAGG...CGG 593

518 ValArgPheValAspGluValAlProLysGlyLeuThrLysIle 532

594 GTTACCTTCGTGAGAGCGTCCAAATACAGCCGACGCAAGATC 638

seq_name: gb_est1:AW691106

seq_documentation_block:

LOCUS AW691106 656 bp mRNA EST 15-JUN-2000

DEFINITION NF04IC08ST1F1000 Developing stem Medicago truncatula cDNA clone

NEF04IC08ST 5', mRNA sequence.

ACCESSION AW691106

VERSION AW691106.1 GI:7565842

KEYWORDS EST.

SOURCE barrel medic.

ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliales; Medicago.

REFERENCE 1 (bases 1 to 656)

He, X.-Z., Shadle, G., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W., May, G.D. and Dixon, R.A.

Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula stem library

UNPUBLISHED (2000)

CONTACT: Dixon RA

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7302

Fax: 580 221 7380

Email: radixon@noble.org

Insert length: 656 Std Error: 0.00

Plate: 041 Row: C Column: 08

Seq primer: TCACACAGAAACGCTATGAC.

```

FEATURES
  source
    Location/Qualifiers
      1. 656
        /organism="Medicago truncatula"
        /db_xref="taxon:3880"
        /clone="MF041C08SR"
        /clone_lib="Developing stem"
        /tissue_type="stem"
        /dev_stage="Pooled developmental"
        /note="Vector: Lambda Zap; Contains a mixture of
        internal stem segments"
BASE COUNT      223 a      121 c      155 g      157 t
ORIGIN
alignment_scores:
  Quality: 504.50      Length: 219
  Ratio: 3.193      Gaps: 2
Percent Similarity: 72.146      Percent Identity: 46.575
alignment_block:
  US-09-581-241-4 x AW691106 ..
Align seg 1/1 to: AW691106 from: 1 to: 656
317 GlyGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgph 333
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
3 GGCGCTGCTCCATCGAATGGAACTGTAACACGCTGTTAAAGATAGGTT 52
333 eAsnLeuProGlyValArgGlnGlyTyrGlyLeuThrGlu..... 346
53 GCCCGAAGCTGACTTGGACAGGATATGGATATGACAGAGCGACGACCAC 102
347 ..ThrThrSerAlaIleIleThrProGluGlyAspAspLysProGly 362
103 TTTCATATAGCTTGGCATTTGCCAAGAAAGAACCATTCAGACAAACCTGGC 152
363 AlaSerGlyLysValValProLeuPheLysAlaLysValIleAspLeuAs 379
153 GCATGTGGCAGTGTGAAGAAACGCTGAGTGAAGTAAGTGGACACAGA 202
379 pThrLysLysThrLeuGlyProAsnArgArgGlyGluValCysValLysG 396
203 AACTGTGCTCCCTTCCAGAAAGAAAGCTGGGAATTTGTTATGAG 252
396 LyrProMetLysMetLysGlyTyrValAspAsnProGluAlaThrArgGlu 412
253 GCACCAAGCTTATGAAGAGATACCTAATGACCTGAGCGCAACAGAGA 302
413 IleIleAspGluGlyTyrLeuHisThrGlyAspIleGlyTyrTyrAs 429
303 ACTATATGACAAAGAGAGATGCTACACAGAGTGACATTGGTTAATTGA 352
429 pGluGluLysHisPhePheIleValAspArgLeuLysSerLeuIleLysT 446
353 CGACGACGAGAGAACTCTTCATTTGTTGATTCATTAAGCAATTGATTAAAT 402
446 YrLysGlyTyrGlnValProProAlaGluLeuGluSerValLeuLeuGln 462
403 ACAAGAGATGCCAAGTAGCTCTGAGCTCGAAGCATTTGATTGATGCT 452
463 HisProAsnIlePheAspAlaGlyValAlaGlyValProAspProIleAl 479
453 CATTCACAAACATTTCTGATGCTGCTGCTGATCCATTTGAAGATGAAGCTGC 502
479 aglyGluLeuProGlyAlaValAlaValLeuLysLysGlyLysSerMetr 496
503 TGGAGAGAGTCCAGTTGCTTTGTTGTAAGATCAAAATGCTTCAAAGTCA 552
496 hGluLysGluValMetAspTyrValAlaSerGlnValSerAsnAlaLys 512
553 GCGAGGATGAATCAAGCAATACATATCCGACAGAGTTGTAATTTTACAG 602
513 ArgLeuArgIlyGlyValAlaArgPheValAspGluValProLysGlyLeuTh 529

```

```

||||: ||| |||:||||: |||:||||: |||:
603 AGAATA...AACAGATTACTTACAGAAACATTTCTTAAGCGCGCTC 649
529 rGlyLys 531
650 AGGAAA 656
seq_name: gb_est1:A1485586
seq_documentation block:
LOCUS      A1485586      625 bp      mRNA      EST      18-MAY-2001
DEFINITION EST243907 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
CLEVED7A3, mRNA sequence.
ACCESSION A1485586
VERSION   A1485586.1 GI:4380957
KEYWORDS  EST.
SOURCE    tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 625)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.,
Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman
,C.L., Niemam,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley
,S.D. and Giovannoni,J.
Generation of ESTs from tomato carpel tissue
Unpublished (1999)
CONTACT: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES
  source
    Location/Qualifiers
      1. 625
        /organism="Lycopersicon esculentum"
        /cultivar="TA496"
        /db_xref="taxon:4081"
        /clone="CLEVED7A3"
        /clone_lib="tomato ovary, TAMU"
        /tissue_type="carpel"
        /dev_stage="5 days pre-anthesis to 5 days post-anthesis"
        /lab_host="XL1-Blue MRF"
        /note="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2:
        XhoI; cLED - Tomato Carpel EST Library. OligodT-primed and
        directionally cloned cDNA in vector Lambda Zap II with 5'
        and 3' ends located at the EcoRI and XhoI sites,
        respectively."
BASE COUNT      197 a      112 c      146 g      169 t      1 others
ORIGIN
alignment_scores:
  Quality: 482.50      Length: 200
  Ratio: 3.260      Gaps: 1
Percent Similarity: 74.000      Percent Identity: 48.500
alignment_block:
  US-09-581-241-4 x A1485586 ..
Align seg 1/1 to: A1485586 from: 1 to: 625
314 IleAlaSerGlyAlaProLeuSerLysGluIleGlyGluAlaValAl 330
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
17 GTATATGCTGGGGCTGCACCATTAAGCAAACTTGAACACACTGTTGC 66
330 aArgArgPheAsnLeuProGlyValArgGlnGlyTyrGlyLeuThrGluT 347
67 AGCCAAATTTCTAATGCTTAAGTGTTCAGAGTTACGAGATGACAGAG 116
347 hrThrSerAlaIleIleIle.....ThrProGluGlyAsp 358
||||: ||| |||:||||: |||:||||: |||:

```

```

117 CCGAGCAGTGTGGCTATGTGCTTGGCATTTGCTAAGAACCCCTCGAA 166
359 AsplysProglYAlaSerGlyLysValProleuphelsAlaLysVa 375
167 ATTAATCTGGAGCATGTGGAGCTGTGTAGAAATGCTCAGATGAAAT 216
375 lIleAspleuAspThrLysThrLeuGlyProAsnArgArgLysLuv 392
217 TGTGGATCCTGATACTGGCAATCTCTCCTAGAACCAATCTGGAGAGA 266
392 aLcysValLysGlyPrometLeuMetLysGlyTyValAspAsnProglu 408
267 TTTGATATAAGAGCGCATCAATTAAGAAGTTACTTAATGATCCAGAG 316
409 AlathrArgLuiIleIleAspGluGluGlyTrpLeuHisThrGlyAsp11 425
317 GCCACTACGGAGCAATAGACAAAGAAAGGTGCTACATACGGCGATAT 366
425 eGlyTyTrpAspGluGluLysHisPhePheIleValAspArgLeuLys 442
367 TGTATATATGTGATCATGATGATGATGATGATGATGATGATGATGAT 416
442 eLleuIleLysTyLysGlyTyValProProAlaGluLeuGluSer 458
417 AATTGATTAATAACAAGATTTCAAGTGGCTCTGCTGAATCTGAAGCC 466
459 ValLeuLeuGlnHisProAsnIlePheAspAlaGlyValAlaGlyValPr 475
467 CTCTCTCCATCAATCATCCCAATTTTCAGATGCTGCTGTTTCATGAA 516
475 oAspProIleAlaGlyLysLysProglYAlaValAlaValLeuLysLysG 492
517 AGACGAGCAAGCAAGAAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 566
492 lYlSerMetThrGluLysGluValMetAspTyValAlaSerGlnVal 508
567 GCTCCCAATCTACTGAAGATCAAGATCAATCATCAACANAGCAGGTG 616
seq_name: gb_est1:BE454671

```

```

seq_documentation_block:
LOCUS BE454671 686 bp mRNA EST 02-MAR-2001
DEFINITION HVSMEN0094P16f Hordeum vulgare 5-45 DAP spike EST library
            HVSDNA0009 (5 to 45 DAP) Hordeum vulgare cDNA clone HVSMEN0094P16f,
            mRNA sequence.
ACCESSION BE454671
VERSION BE454671.2 GI:13189489
KEYWORDS EST.
SOURCE barley.
ORGANISM Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
            ; Triticeae; Hordeum.
            1 (bases 1 to 686)
Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu
Y., Anderson,H., Dale,J., Henry,D., Kernodle,S., Palmer,M., Rambo
T., Saski,C., Schwartzbeck,J., Simmons,J., Choi,D.W., Main,D. and
Wood,T.
Development of a genetically and physically anchored EST resource
for barley genomics
Unpublished (2000)
Contact: Wing RA
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: AATTAACTCATCACTAAAGG
High quality sequence stop: 653.
Location/Qualifiers
source 1. 686

```

```

/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSMEN0094P16f"
/clone_lib="Hordeum vulgare 5-45 DAP spike EST library
HVSDNA0009 (5 to 45 DAP)"
/tissue_type="5-45 DAP spike"
/lab_host="SOLR"
/note="Vector: LambdaZAP, Site_1: EcoRI, Site_2: XhoI. For
more details on library preparation and sequence analysis
see http://www.genome.clemson.edu/projects/barley/ To
order a clone see http://www.genome.clemson.edu/orders"
BASE COUNT 156 a 212 c 203 g 115 t
ORIGIN
alignment_scores:
Quality: 482.00 Length: 223
Ratio: 2.957 Gaps: 2
Percent Similarity: 73.094 Percent Identity: 43.049
alignment_block:
US-09-581-241-4 x BE454671 ..
Align seg 1/1 to: BE454671 from: 1 to: 686

```

```

314 lIleAspGlyAlaProleuSerLysGluIleGlyGluAlaValAl 330
17 GTCATGTCCGGCGCGCCATGCGCAAGAGCTGACGACGCTTCAT 66
330 aArgArgPheAsnLeuProglYAlaArgGlnGlyTyValLeuThrLut 347
67 GGCCAAGATCCCAAGCGCGTCTCGGGCAGGGGTACGGATGACGAGG 116
347 hTrhSerAlaIleIleIleThrProGluGlyAspAspLysPro..... 361
117 CCGGCGCGTCTGCGCATGTGCTGCGCTGCGTCCGCAAGAGCCCTTCAG 166
362 .....GlyAlaSerGlyLysValAlaProleuphelsAlaLysVa 375
167 GTCAAGTCCGGGTCTGCGGCGCATGCTGCGCAAGCCGGGCTCAAGAT 216
375 lIleAspleuAspThrLysThrLeuGlyProAsnArgArgLysLuv 392
217 CGTCGACCCCGACACGGCGCTCTGCGCGGAGAACACGCGCGGAGAGA 266
392 aLcysValLysGlyPrometLeuMetLysGlyTyValAspAsnProglu 408
267 TCTGCATCCGGGGGAGCAGATCATCAAGAGTTACCTTAACGACCCAGAA 316
409 AlathrArgLuiIleIleAspGluGluGlyTrpLeuHisThrGlyAsp11 425
317 TCACACCAAGAACACCATGACAGAGATGCTGCTGCGACACCGAGACAT 366
425 eGlyTyTrpAspGluGluLysHisPhePheIleValAspArgLeuLys 442
367 CGGACATCTTCTGGCGAAGTCCCATGCTCTTATGCGGATCGAAG 416
442 eLleuIleLysTyLysGlyTyValProProAlaGluLeuGluSer 458
417 ACATCATCAAGTACAAGGCTTCCAGGTGCGCGGCGGACGACTCGAGGCC 466
459 ValLeuLeuGlnHisProAsnIlePheAspAlaGlyValAlaGlyValPr 475
467 CTCTCATCACCCACCGGAGATCAAGACGCGCGCTGATGCTGAA 516
475 oAspProIleAlaGlyLysLysProglYAlaValAlaValLeuLysLysG 492
517 GGACGATCTTCTGCGGAGATCCCATGCTCTTATGCGGATCGAAG 566
492 lYlSerMetThrGluLysGluValMetAspTyValAlaSerGlnVal 508
567 GCTTGAATTCACGAGATGATCAAAAAAATTCGTCCGAAAGAGTT 616

```


VERSION BF064865.2 GI:13266207
 KEYWORDS EST.
 SOURCE barley.
 ORGANISM Hordeum vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
 ; Triticeae; Hordeum.
 REFERENCE 1 (bases 1 to 806)
 AUTHORS Ming, R., Close, T.J., Kleinholz, A., Wise, R., Begum, D., Frisch, D., Yu
 , Y., Anderson, H., Dale, J., Henry, D., Kernodle, S., Palmer, M., Rambo
 , T., Saski, C., Schwartzbeck, J., Simmons, J., Choi, D.W., Math, D. and
 Wood, T.
 TITLE Development of a genetically and physically anchored EST resource
 for barley genomics
 JOURNAL Unpublished (2000)
 COMMENT On Oct 17, 2000 this sequence version replaced gi:10841504.
 CONTACT: Ming RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: twing@clemson.edu
 Seq primer: AATTAACCTCCTCAAGAGG
 High quality sequence stop: 759.
 FEATURES
 source
 location/Qualifiers
 1..806
 /organism="Hordeum vulgare"
 /cultivar="C116151 (M1a6)"
 /db_xref="taxon:4513"
 /clone="HV_CEP0021D12f"
 /clone_lib="Hordeum vulgare seedling green leaf EST
 library HYCDNA0005 (Erysiphe infected & control)"
 /tissue_type="Seedling green leaf"
 /lab_host="SOLR"
 /note="Vector: lambdaZAP; Site_1: EcoRI; site_2: XhoI; For
 more details on library preparation and sequence analysis
 see http://www.genome.clemson.edu/projects/barley/ To
 order a clone see http://www.genome.clemson.edu/orders"
 BASE COUNT 185 a 201 c 239 g 181 t
 ORIGIN
 alignment_scores:
 Quality: 479.00 Length: 211
 Ratio: 3.051 Gaps: 3
 Percent Similarity: 74.408 Percent Identity: 45.972
 alignment_block:
 US-09-581-241-4 x BF064865 ..
 Align seg 1/1 to: BF064865 from: 1 to: 806
 340 GlnGlyTyrGlyLeuThrGluThrThrSerAlaIleIleIleThrProG1 356
 ||||||||||||||||||||||||||||||||||||||||
 3 CAGGGGAGGCGGTGACCGAGACGACGCGCGGTGCGCCACGACG 52
 356 uGlyAspAsp.....LysProGlyAlaSerGlyGlyValValProLeup 371
 ||||||||||||||||||||||||||||||||||||||||
 53 GGGCGGAGAGAGCGCGGTGACGAGCGCGCGGTGCGCCACACA 102
 371 heLysAlaLysValIleAspLeuAspThrLysLysThrLeuGlyProAsn 387
 ||||||||||||||||||||||||||||||||||||||||
 103 CCGAGGCGCAAGATGCTGACCGGAGACGAGCGCGGTGCGGAGAAC 152
 388 ArgArgGlyGlyValAlcysValLysGlyProMetLeuMetLysGlyTyrVa 404
 ||||||||||||||||||||||||||||||||||||||||
 153 CGACACGGGAGGCTGTGATCAGGAGGACCTACGTATGAAGATACCT 202
 404 LAspAsnProGlyAlaIleThrArgGluIleIleAspGluGluGlyTyrLeuH 421
 ||||||||||||||||||||||||||||||||||||||||
 203 CAAGAACACAGAGGCAACACATGCAAGGTGACCGCGCGGATGGCTCA 252

421 IsThrGlyAspIleGlyTyrTyrAspGluGluLysHisPheIleVal 437
 ||||||||||||||||||||||||||||||||||||||||
 253 AGACCGGTGACCTATGCTATACATACGACGAGTGGTATCTCTGCTG 302
 438 AspArgLeuLysSerLeuIleLysTyrLysGlyTyrGlnValProAl 454
 ||||||||||||||||||||||||||||||||||||||||
 303 GACCGCTGAAAGAGTGTATCAATATACAAAGGATATCAGTGCCTCAGC 352
 454 agLLeuGluSerValLeuLeuGlnHisProAsnIlePheAspAlaGly 471
 ||||||||||||||||||||||||||||||||||||||||
 353 AGAGTGTGAACCTCTTGTGCTGACCCATCCAGAGTTTCCGATGCTG 402
 471 aLaIaGlyValProAspProIleAlaGlyGluLeuProGlyAlaVal 487
 ||||||||||||||||||||||||||||||||||||||||
 403 TTATTCCTTCCACAGACGAGGAGTGGTCACTCCGATGCGCTATGTC 452
 488 ValLeuLysLysGlyLysSerMetThrGluLysGluValMetAspTyrVa 504
 ||||||||||||||||||||||||||||||||||||||||
 453 GTGAGGAAGAAAGGAGCAATTTGTCAAGCGCAGAGGTGATGAGATTGT 502
 504 lAlaSerGlnValSerAsnAlaLysArgLeuArgGlyGlyValAlaArgPheV 521
 ||||||||||||||||||||||||||||||||||||||||
 503 GCGCAACAGGTAGACCTTACAAAGAGGTGAGG...AAGTGCGCTTGC 549
 521 aLAspGluValProLysGlyLeuThrGlyLysIleAspGlyAlaIle 537
 ||||||||||||||||||||||||||||||||||||||||
 550 TGACAGACATCCCAAGATGCGTGTGCAAGATACAGAGGAGGAGGAGGAG 593
 538 ArgGluIleLeuLysLysProValAlaLysMet 548
 ||||||||||||||||||||||||||||||||||||||||
 594 ...GATCTTATCAAGCTCGCAGCCTCCACAGCTG 623
 seq_name: gb_est2:BE820931
 seq_documentation_block:
 LOCUS BE820931 772 bp mRNA 24-MAY-2001
 DEFINITION GW70001A20B2 Gm-r1070 Glycine max cDNA clone Gm-r1070-5044 3',
 mRNA sequence.
 ACCESSION BE820931
 VERSION BE820931.1 GI:10253165
 KEYWORDS EST.
 ORGANISM soybean.
 Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseolae;
 Glycine.
 REFERENCE 1 (bases 1 to 772)
 AUTHORS Vockin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V.,
 Expelling, J., Raph, C., Shoop, E., Pardinas, J., Liu, L. and Lewin, H.
 TITLE A Functional Genomics Program for Soybean (NSF 9872565)
 JOURNAL Unpublished (1999)
 COMMENT Other-ESTs: AM458737 corresponding to Gm-cl0106-6629 (5')
 Contact: Vockin, L.O., PI, A Functional Genomics Program for
 Soybean (NSF 9872565)
 Lewin, H. A., Director, Keck Center for Comparative and Functional
 Genomics
 University of Illinois
 Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
 Tel: (217) 244-6147
 Fax: (217) 333-4582
 Email: l-vockin@uiuc.edu
 This clone is available through: Genome Systems, Inc. 4633 World
 Parkway Circle St. Louis, Missouri 63134. For further information
 call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
 427-3324 or contact: clones@genomesystems.com or info@genome
 systems.com web site: www.genomesystems.com
 Seq primer: 5'-TTTTTTTTTTTTTTT(A/C/G)-3'.
 FEATURES
 source
 location/Qualifiers
 1..772
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="Gm-r1070-5044"

US-09-581-241-4 x AV558307 ..

Align seg 1/1 to: AV558307 From: 1 to: 604

```

304 AsPLysTYrAspLeuSerAsnLeuValGluIleAlaSerGlyAlaLys 320
      :::::::::::::::::::::::::::: ::::::::::::::
3  AATCTTATGATCTCTCCCTCCGTTAGATTGCTTCTCCGTCAGCTCC 52
      :::::::::::::: ::::::::::::::
320 oLeuSerLysGluIleGlyIleAlaValAlaArgArpHeAsnLeuProG 337
      :::::::::::::: ::::::::::::::
53  TTAGGTAAAGAACTCAAGATAGTCTGTGCTCCCTCCCAAGCCA 102
      :::::::::::::: ::::::::::::::
337  LYValArgInclYTYrGlyLeuThrGluThrSerAlaIleIle 353
      :::::::::::::: ::::::::::::::
103  TCCTTGCCAGGATATGATGATACGAGCAGCAGCAGTGTGCAATG 152
      :::::::::::::: ::::::::::::::
354  Thr.....ProGluGlyAspAspLysProG 362
      ::::::::::::::
153  AGCCTTGCGTTCCTAAAGACCGATACG.....ACAAATACAG 193
      ::::::::::::::
362  yAlaSerGlyLysValAlaProLeuPheLysAlaLysValIleAspLeuA 379
      :::::::::::::: ::::::::::::::
194  TTCATGTGGAGTGTGTCGGAACGACAGCTTAAGTGTTCACCTTG 243
      :::::::::::::: ::::::::::::::
379  sPThLysLysThrLeuGlyProAsnArgArgGlyIleValCysValLys 395
      :::::::::::::: ::::::::::::::
244  AGACACGTCCTCTCTGTGATCAACCAACCTGTGATTTGTATCCGC 293
      ::::::::::::::
396  GLYPromLeuMetLysGlyTYrValAspAsnProGluAlaThrArgG 412
      ::::::::::::::
294  GGTCAACAGATCATGAAGAGTACTTGAACGATCCAGAAAGCCACTTCAGC 343
      ::::::::::::::
412  uIleIleAspGluGluGlyTYrPheuHisThrGlyAspIleGlyTYrA 429
      ::::::::::::::
344  AACATAGACAGCAAGAGTTGGCTTCACACAGACAGCATTTGATATG 393
      ::::::::::::::
429  spGluGluLysHisPhePheIleValAspArgLeuLysSerLeuIleLys 445
      ::::::::::::::
394  ATGAAGATGATGAGATTTCATTTGATCGACTTAAAGAACTCATCAAG 443
      ::::::::::::::
446  TYrLysGlyTYrGluValProProAlaGluLeuGluSerValLeuLeuG 462
      ::::::::::::::
444  TTCAAAGGCTTCAGGCTCCCTCCAGCTGAGCTAGAGAGTTTCTATCA 493
      ::::::::::::::
462  nHisProAsnIlePheAspAlaGlyValAlaGlyValProAspProIleA 479
      ::::::::::::::
494  TCACCATTCATTCATTCGCGATGCGCTGTGTCGCGAAATGATGAAGTCG 543
      ::::::::::::::
479  IaglyGluLeuProGlyAlaValAlaValLeuLysGlyLysSerMet 495
      ::::::::::::::
544  CTGGGGAAGTCCGCGTTCGTCGTGCGATCAAAATGGAATGATATC 593
      ::::::::::::::
496  ThrGluLys 498
      ::::::::::::::
594  ACTGAAGAA 602

```


	Strc	Orig	Len	Docum
gb_gss:BGN01FCA	+625.00	1200.41	975	AL11659 Anopheles gambiae GSS
gb_est2:BG454563	+552.00	1058.48	882	BG445563 CA_PAO028K14f Gossypid
gb_est1:BG365528	+541.00	1034.95	1066	BEO365528 MP01A02 MP Meemibryara
gb_est1:AU014129	-513.00	983.93	741	AU014129 unpublished dltic
gb_est2:C83857	-512.00	983.12	664	C83857 Dicyostellum dltic
gb_est2:BF49Z6267	-507.50	974.01	1.3e-45	BF49Z6267 AT01J136_3prime AT Droso
gb_est1:AI1846799	+505.50	970.34	6.9e-45	AI1846799 EST545121 tomato ovary1
gb_est2:BG414540	+505.50	967.48	1.0e-44	BG414540 HYSMEM0020N3f Hordenu
gb_est1:AM691106	+504.50	968.56	2.7e-45	AM691106 NF04C108STRT1P1000 Develo
gb_est1:A1485566	+482.50	925.98	8.1e-45	A1485566 EST243907 tomato ovary1
gb_est1:BB454671	+482.00	925.02	2.6e-42	BB454671 HYSMEM0094p16f Hordenu
gb_est2:BF623201	+481.50	921.21	3.8e-42	BF623201 HYSMEM0015K22f Hordenu
gb_est2:BF063485	+479.00	916.45	7.0e-42	BF063485 HV_CEB0021D12f Hordenu
gb_est2:BBE820937	-477.50	913.97	9.6e-42	BBE820931 GM700013KA2082 Gm-tr1077H2
gb_est1:AVS580391	+474.50	910.67	1.5e-41	AVS58307 AV058307 Arabidopsis t
gb_est1:AW329421	+472.50	900.99	2.3e-41	AW329421 N200656e rootphos(-) N
gb_est1:AU214473	-472.50	906.19	2.6e-41	AU214473 AU0214473 unpublished c
gb_est1:BE555223	+469.50	901.99	4.5e-41	BE555223 spB7440_y1 Gm-cl045 GSS
gb_gss:ONS01MDU	-468.00	892.89	1.4e-40	AL150803 Anopheles gambiae GSS
gb_gss:ONS01IBD	-466.00	891.49	1.7e-40	AL149424 Anopheles gambiae GSS
gb_est2:BI311229	+465.50	891.60	1.7e-40	BI311229 EST531297f GRSD Medice
gb_est1:AW621420	+464.50	891.47	1.7e-40	AW621420 EST312218 tomato roo
gb_est1:AA979869	+464.00	889.64	2.8e-40	AA979869 MEST2-F3_TW1412_Seq Iri
gb_est1:AV680327	-463.50	889.95	2.1e-40	AV680327 AV680327 Norci Satoh unc
gb_est1:AU2193664	-458.50	879.33	8.2e-40	AU2193684 AU219684 unpublished c
gb_est2:BG506544	+456.00	875.09	1.4e-39	BG506544 EST440621 CYN Mediaqsg
gb_est1:AW218471	+454.50	872.12	2.1e-39	AW218471 EST303654 tomato tradit
gb_est1:AW560940	+453.50	868.69	3.2e-39	AW560940 EST315988 DSIR Medicaea
gb_est2:BF200141	+451.50	862.84	6.4e-39	BF200141 WHE2252_E02_I04E2 Trifl
gb_est1:AM393904	+450.00	862.61	5.8e-39	AM392094 N200303e rootphos(-)
gb_est2:BF634445	+450.00	862.36	7.2e-39	BF634445 NF056209PTRT1076 Drouca
gb_est2:BI309413	+447.50	857.11	1.4e-38	BI309413 EST530823 GRPD Medicaea
gb_est1:BB572743	+445.50	853.20	2.3e-38	BB597274 P11-70_CO5_91 AU02 Paeca
gb_est1:AJ284477	-445.00	850.24	1.9e-38	AJ284477 4A3B-AAx-B-O4 F Anophh
gb_est1:AL393983	-443.00	855.97	3.4e-38	AL393982 MBCE57HO9F1 MtBC Medici
gb_est1:AW776080	+443.00	848.61	4.2e-38	AW776080 EST335145 DSLI Medicaea
gb_est1:AI1488821	+438.00	840.71	1.2e-37	AI1488821 EST247160 tomato ovary1
gb_est1:BB330067	+437.50	838.90	1.5e-37	BB330067 sc73a05_y1 Gm-cl040 G
gb_est2:BB080370	+435.50	833.63	2.8e-37	BB080370 sr75se03_y1 Gm-cl051 G
gb_est2:BI8435803	+434.50	833.35	2.6e-37	BI842803 sah7t2i03_y1 Gm-cl049 G

234 CCTAAGGAGTTGAGTTGACGCATCAGAGCATTATTGCTACGATCGCACA 283

```

223 sAlaArGAsPProIleTyrgLyAsnGlnValSerProGlyThrAlaIleL 240
284 CTCGAAGAAGTGGGCTAAGCTGCTGAGAGCTACCCGATGAGCTGTAGCAC 333
240 eUrThValValProPheHisHis.....GlyPheGlyMetPheThr 233
334 TAGCACTGACTCTCTGTTCCAGTCGCGCCGCTGCTGCGGCTG..... 377
254 ThrLeuGlyThrLeuThrCysGlyPheArgLLeValMetLeuThrLysPh 270
378 ...CTGAATATGTTACCAACAATTCGCCGTGCTGTGCTGATGCCAGATT 424
270 eAspGluGluThrPheLeuLysThrLeuGlnAspTyrlYsCysSerSery 287
425 TCATCCACACCTGTTCTTACGTATGTAAGAATATCGGGTAATCTAA 474
287 aLlLeuValProThrLeuPheAlaIleLeuAsnArgSerGlyLeuLeu 303
475 TCACACTGCTACCTCCGCTGATGCTTTCTTGCCCAACATCCCATGGTG 524
304 AspLysTyAsPLeuSerAsnLeuValIleGlyLeuAlaSerGlyValAlaP 320
525 GACACTAGCATCTCTCTCGCTGATGAGCGTGTCTGTGCGCGGCTCC 574
320 oLeuSerLysGluIleGlyGlnAlaValAlaArgArgPheAsnLeuProG 337
575 GTTGAGCAAGAAGATCAGATCAGTGGCGGAGCGACGTGGGTAGCTT 624
337 lYValArgGlnGlyTyrgLyLeuThrGluThrThrSerAlaIleIle 353
625 TCATTTCGAGGAGTACGGCATGATGAAGCAACCTCGTGTGCTGATG 674
354 ThrProGluGlyAspAspLysProGlyAlaSerGlyLysValValProL 370
675 CAGGACGAGTTGAGAACAAAGGGGACAGCTGGGAAAGTGGCGCATGGG 724
370 uPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGlyPro 387
725 CCAGTGGGTGAAGGTGATGACCCGAGGAGCGGCGCTGCGCGACGCA 774
387 sAlaArgGlyGlyValLysValLysGlyProMetLeuMetLysGlyTyrl 403
775 ACCAGCGGGGTAGCTGCTTTAAGGCTCATGTGATGAAGAAGATAT 824
404 ValAspAsnProGluAlaThrArgLLeIleAspGluGluGlyTrpLe 420
825 GTGGCGAAGGACAGGCC.....ATCGATGCTGACGCTGGCT 862
420 uHisThrGlyAspIleGlyTyrlYrAspGluGlyLysHisPhePheIle 437
863 GCACACGGGCGACATTGCGTACTACATGATGATCAGAGTTTTCATCG 912
437 aLAspArgLeuLysSerLeuIleLysTyrlYsGlyTyrgLValProPro 453
913 TCGACCGGATCAAGAGCTGATCAAGAGGCTTCCARGTACCGCGG 962
454 AlaGluLeuGlu 457
963 GCCGAGCTCGAA 974

seq_name: gb_est2:BG445463

seq_documentation_block:
LOCUS      BG445463          882 bp      mRNA          EST          15-MAR-2001
DEFINITION GA_Ea0028K14f Gossypium arboreum 7-10 dpa fiber library Gossypium
arboresum cDNA clone GA_Ea0028K14f, mRNA sequence.
ACCESSION  BG445463
VERSION     BG445463.1  GI:13355211
KEYWORDS   EST.
SOURCE      Gossypium arboreum.
ORGANISM   Gossypium arboreum; Streptophyta; Embryophyta; Tracheophyta;
Eukaryota; Viridiplantae;

```

```

REFERENCE  1 (bases 1 to 882)
AUTHORS   Wing,R.A., Fritsch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
           D., Wood,T.C., Leslie,A. and Wilkins,T.A.
TITLE      An integrated analysis of the genetics, development, and evolution
           of the cotton fiber
JOURNAL    Unpublished (2000)
COMMENT    Contact: Wing RA
           Clemson University Genomics Institute
           Clemson University
           100 Jordan Hall, Clemson, SC 29634, USA
           Tel: 864 656 7288
           Fax: 864 656 4293
           Email: rwing@clemson.edu
           Seq primer: TAATACGACTCACTATAGG
           High quality sequence stop: 733.
           Location/Qualifiers
             1..882
               /organism="Gossypium arboreum"
               /strain="AKA"
               /cultivar="8400"
               /db_xref="taxon:29729"
               /clone="GA_Ea0028K14f"
               /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
               /tissue_type="Fibers isolated from bolls harvested 7-10
               dpa"
               /lab_host="E. coli"
               /note="vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 250 a 173 c 232 g 226 t 1 others
ORIGIN

alignment_scores:
  Quality: 552.00      Length: 277
  Ratio: 2.816        Gaps: 4
  Percent Similarity: 70.758  Percent Identity: 41.155

alignment_block:
US-09-581-241-6 x BG445463 ..

Align seg 1/1 to: BG445463 from: 1 to: 882
263 ArgIleValMetLeuThrLysPheAspGluGluThrPheLeuLysThrLe 279
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6 CGAGTTTATCATCGAAGATTCGACATCGCGCTTGTGAGCTAAT 55
279 uGlnAspTyrlYsCysSerSerValIleLeuValProThrLeuPheAlaI 296
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
56 ACAGAAATACAAAGTACCATTTGCTCGATTGCGCCAGATGTTTGG 105
296 lLeuAsnArgSerGluLeuLeuAspLysTyrlYsPheSerAsnLeuVal 312
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 CCATCGCTAAGTCATCGGAAGCAAAATTTACGATTTGTCATCGCGAGG 155
313 GluIleAlaSerGlyGlyAlaProLeuSerLysGluIleGlyGlnAlaVa 329
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
156 AAGTTGAAGTCGCGTCCGCTCGCTGGGTCAAGAGCTTGAAGATCTGT 205
329 lAlaArgArgPheAsnLeuProGlyValArg.....GlnGlyTyrgLy 344
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
206 AAAAGTCAAGTTT.....CTGTGTCCAAAATTTGACAGAGGTTATGAA 249
344 eUrThGluThrThrSerAlaIleIleIle.....ThrPro 355
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
250 TCAGACAAGCTGGACCAAGTTTTCAGCAATGTGTTGGGATTTGCCAAGAA 299
356 GluGlyAspAspLysProGlyAlaSerGlyLysValValProLeuPheLy 372
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
300 CCTTTGAATGAATCCGAACTTGTGGAGCTGTTGTAAGAACCGGA 349
372 sAlaLysValIleAspLeuAspThrLysLysThrLeuGlyProAsnArgA 389
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

	seq_name:	gb_estl:BE036528	
	seq_documentation_block:		
LOCUS	BE036528	1066 bp	mRNA
DEFINITION	MP01A02 MP Mesembryanthemum crystallinum cDNA 5' similar to	EST	07-JUN-2000
ACCESSION	BE036528		
VERSION	BE036528.1	GI:8331537	
KEYWORDS	EST.		
SOURCE	common ice plant.		
ORGANISM	Mesembryanthemum crystallinum		
	Eukaryota; Vitidiplanteae; Streptophyta; Embryophyta; Tracheop		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
	Caryophyllidae; Caryophyllales; Aizoaceae; Mesembryanthemum.		
REFERENCE	1 (bases 1 to 1066)		
AUTHORS	Bohnert,H.J., Borchert,C., Brazille,S., Brooks,D., Eaton,M., H., Kawasaki,S., McColough,A., Michalowski,C.B., Palacio,C. Sara,G., Wheeler,M. and Zepeda,G.R		
TITLE	Functional Genomics of Plant Stress Tolerance		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Michalowski,C.B. University of Arizona Bio Sciences West room 513, Tucson, AZ 85721, USA Tel: 520-621-7982 Fax: 520-621-1697 Email: chmeh.arizona.edu An open reading frame exists. Insert Length: 1 Std Error: 0.00. Location/Qualifiers 1..1066 /organism="Mesembryanthemum crystallinum" /db_xref="taxon:3544" /clone_lib="Mp" /issue_type="apical meristem and leaf primordia" /dev_stage="6 weeks"		

```

BASE COUNT      265 a      220 c      279 g      294 t      8 others
ORIGIN

alignment_scores:
    quality: 541.00      length: 295
    Ratio: 2.678      Gaps: 5
    Percent Similarity: 68.475      Percent Identity: 40.339

alignment_block:
US-09-581-241-6 x BE036528 ..

Align seg 1/1 to: BE036528 from: 1 to: 1066

218 VALTHRAAPHNE.....SEHISALARGASPROLETTYRGLYSNGI 232
      ::::::::::: ||||| :::::::::::
4 ATTACGGCGTATAGTAGTGCATATTAATAATCTCG..... 39

232 nvalserproglythrallaileuthvalvalprophenishisglyp 249
40 .....CCAAATGGGCGCTGTTCTCCGCGTTTCATGAGTCT 76

249 heglymethethrthrleuglytyrleuthcysglypheatgileval 265
      |||||::: ||| ::::: :::::
77 TGGGTTCCTTCATGTGCTCTCGAGAGCGGTGCTGGCCGAACATAGT 126

266 Metleuthrlysrphasprogluuthrphenleuthrleuglnaspty 282
127 GTAATGGCAAGTGTGATTTGAGAAAGATGTGGAGGCTGGAGAAAT 176

282 rlyscyserservalilleuvalprothrphenallaileuasn 299
      ::::: ::::: ||| ||| :::::
177 TCGGATTACATCATTCGCGTTCGCCGCCATTTGGTGTGCGCATGCAA 226

299 rgserserleuenuhsprlytyraspleusesarnleuvalgileala 315
227 AATCGAGGCTTACATATAGTACATCTCAGCAGCGTGGAAAGTCTGCT 276

315 SerglylyalaProleuserlyscuileglylualavalalaarg 332
277 TGNAGTGGCGCTCTCTTGGGAAGAGAGTTCACAGCGTTTCGAGCTG 336

332 gpleasleuProglyvalarglnglytyrlyleuthrcluthrhrs 349
      |||| ::::: ||||| ||||| ||||| |||||
332 GTTCCGGAACGTGGAGATAGCACAGGATATGAGCTGACAGACACAG 376

349 eralallellellethr.....ProgluglyaspasplyProglyala 363
377 GCGCGGCTCCAAAGACGTTAGGCGCTGAGAAATCGAAACGATATGCT 426

364 SerglylyvalvalProleuphelysalalyvalilleasleuaspth 380
      ||||| ::::: ||||| ::::: ||||| :::::
427 GTTGGCGCGCTTACAGAGCTTGTGGAAACAAAGATAGTTGATCCGTC 476

380 rlyslsyrhrleuuglyProasnargarglyglvalaCysvallysglyp 397
477 CGAGAGGCGCTTACCACTCGTCAAGAGGGGAGCTGTGGCTGGCTGGC 526

397 romelumeulleysglytyrvalaspsnProglualathrargjulle 413
      || ::::::::::: ||||| ::::: ||||| :::::
527 CAACCATATGAAGAAGCTATGTGGAGATGAAGATGCCACGTGCACACT 576

414 Ileasplugluythrleuhsrhrclvaspilleglytyrtyraspgl 430
577 NTCCATCCANATGATGAGGTGCAAGCTGTGATCTTGTATTTTGACTC 626

430 ugluysrshaspherhilevalasparleuhserserleuleystyrl 447
627 TGATGGCTTTTCATATGTTGATGCATTTAAAGAAATGATTAAGTACA 676

447 ysglytyrclnvalProproalaglucluseValleuclnhs 465
      ||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

677 AGGCATATCATNGTTCCGCCAGCTGAGTTGAGGCGTTTACTACATATCCAT 726
464 PROAS11IleheaspalaaglyValAlaIaglyValProAspProIlealag1 480
|||||:||||| ||||| ||||| ||||| |||||
727 CCAGATATGTTGCTGATGCTGCTGTTGTCCTATCCGATATGACGAAGCG 776
480 yglueupProGlyAlaValAlaValAlaValLeuylslysgly.LySserMetThr 496
|||||:||||| ||| ||||| ||||| |||||
777 ACAAAATACCATATGCATTTGTGCTC...AGCAATNCAGAAACCATTTTC 823
497 GLuLySGluValMetAspTyrValAlaIaSerGln 507
|||||:||||| ||||| ||||| ||||| |||||
824 TGCTTCACGTCATGATGATATGAGCTAACAGG 856

seq_name: gb_est1:AU214129

seq_documentation_block:
LOCUS AU214129 741 bp mRNA EST 17-JUL-2001
DEFINITION AU214129 unpublished oligo-capped cDNA library, stage L2
ACCESSION AU214129
VERSION AU214129
KEYWORDS AU214129.1 GI:14852286
SOURCE EST.
ORGANISM Caenorhabditis elegans.
Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
1 (bases 1 to 741)
REFERENCE Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
AUTHORS and Sugano,S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished (2001)
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
FEATURES
Source
1..741
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk809f05"
/cclone_lib="unpublished oligo-capped cDNA library, stage
L2"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L2"
BASE COUNT 216 a 158 c 125 g 241 t 1 others
ORIGIN

Alignment_scores:
Quality: 513.00 Length: 223
Ratio: 3.090 Gaps: 5
Percent Similarity: 74.439 Percent Identity: 47.982

alignment_block:
US-09-581-241-6 x AU214129/rev ..

Align seg 1/1 to reverse of: AU214129 from: 1 to: 741

320 ProlueserLySGluIleGlyGluAlaValAlaIaArgArgPhe...AsnIe 335
||| :|||:||||| ||| :|||:||||| |||||
739 CCAGCTGGAAGAGATTTGATTGACAGATTTGAAGAGAGATATACCAATTT 690
335 uProGlyValArgGlnGlyTyrGlyLeuThrGlnuThrThrSerAlaIle1 352
| :|||:||||| ||||| ||||| ||||| |||||
689 GAAATACATCCCAACAGGATATGGAATGACGGAATGTTCAATAGCGTAGTC 640
352 lelleThrProGluIuLyAspAspLyPro...GlyAlaSerGlyLySVal 367

```

LOCUS C83857 664 bp mRNA EST 28-APR-1999
DEFINITION C83857 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
discoideum cDNA clone SSA391, mRNA sequence.
ACCESSION J05644
VERSION C83857
KEYWORDS EST.

SOURCE Dictyostelium discoideum.
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium,
1 (bases 1 to 664)
Moriya,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
Williams,R., Mita,A.M., Pl.M., Sato,T., Takemoto,K., Yasukawa,H.,
Yoshino,J., Maeda,M., Takeuchi,I., Ochiai.H. and Tanaka.Y.
The Dictyostelium developmental CDNA project: generation and
analysis of expressed sequence tags from the first-finger stage of
development
JOURNAL DNA Res. 5 (6), 335-340 (1998)

MEDLINE 99156227
COMMENT Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
3-3-10 Ten-nodaI, Tsukuba, Ibaraki 305, Japan
Email: d402hu@sakura.cc.tsukuba.ac.jp

PROJECT = 'Dictyostelium discoideum cDNA project in Japan'
POLYA-NO.

FEATURES
source location/Qualifiers
1..664
/organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SSA391"
/clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
/dev_stage="slug"

BASE COUNT 234 a 105 c 124 g 201 t
ORIGIN

alignment_scores:
Quality: 512.00 Length: 200
Ratio: 3.261 Gaps: 0
Percent Similarity: 78.500 Percent Identity: 47.500
alignment_block:
US-09-581-241-6 x C83857 ..

Align seg 1/1 to: C83857 from: 1 to: 664

```

340 GlnGlyTyrGlyLeuThrGluThrThrSerAlaIleIleIleThrProG1 356
|||||
2 CAAGTTATGTTGCTACCGAAGCTTACCGAAGCTTGTGTTATTCACAG 51
|||||
356 uGlyAspAspLysProGlyAlaSerGlyLysValProLeuPheLysA 373
|||||
52 TGGATTGATTAAATCTGTTCCGCTGATTCCTTACCAATCAATTCG 101
|||||
373 lAlysValIleAspLeuAspThrLysThrLysThrLysProAsnArg 389
|||||
102 CTAAATATCTCTCCAGAGCTGCTGAGAAATTTAGGTATGGGAAAG 151
|||||
390 GlyLysValCysValLysGlyProMetLeuMetLysGlyTyrValAs 406
|||||
152 GGTAAATCTGTATTAAGGCCAATGTATGTGGTTATTTATACAA 201
|||||
406 nProGluAlaThrArgLuiIleIleAspGluGluGlyTrpLeuHisThr 423
|||||
202 TGAAGAGCAACCAATGATCATAGATAAAGATGATCTTAAAAACTG 251
|||||
423 lYAspIleGlyTyrTyrAspGluGluLysHisPheIleValAsnArg 439
|||||
252 GTGATTGTTACGTCGATGAGATGCTTATTTACTTATCGTCATAGA 301
|||||
440 LeuLysSerLeuIleLysTyrLysGlyTyrGluValProProAlaGlu 456
|||||
302 TCAGAGGATTCATCAATGTAAGGTTCCAGTACTCTCTGCGAATT 351
|||||
456 uGluSerValLeuLeuGlnHisProAsnIlePheAspAlaGlyValA 473
|||||
352 GGAGGCAATTACTATATCTCATCCAAAGTTGCAGATGCTGTAGTAG 401
|||||
473 lYValProAspProIleAlaGlyLysLeuProGlyAlaValValLeu 489
|||||
402 GTCTTTCAAAAGGTGATATGGGTACAGTACCAAGGTTTCGTTGTT 451
|||||
490 LysLysGlyLysSerMetThrGluLysGluValMetAspTyrValAla 506
|||||
452 AAACAAATGAACTCTAACCGAAAGAACTTATGATGGGCCATCC 501
|||||
506 rGlnValSerAsnAlaLysArgLeuArgGlyGlyValAlaPheValAs 523
|||||
502 AAAAATTGCAAAATTATTAACATTTCAAGAGGTGATCTTTTCATAC 551
|||||
523 lValAlaProLysGlyLeuThrGlyLysIleAspGlyLysAlaIleArg 539
|||||
552 CAATTTCCTAAATCTGCAACCGTAACTATTACGTAATAAATCTTAA 601
|||||

```

seq_name: gb_est2:BF492687

seq documentation_block:

LOCUS BF492687 683 bp mRNA EST 19-APR-2001

DEFINITION AT01136.3prime AT Drosophila melanogaster adult testes POTB7
Drosophila melanogaster cDNA clone AT01136 3 similar to CG6178:
FBan0006178 'enzyme' located on: 3R 95E1-95E1: 04/07/2001, mRNA
sequence.

ACCESSION BF492687
VERSION BF492687.1 GI:11575988
KEYWORDS EST.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 683)
Stapleton,M., Brokstein,P., Hong,L., Agbayan,A., Baxter,E., Berman
,B., Carlson,J., Champe,M., Chavez,C., Chew,M., Dorsett,V., Farfan
,D., Frise,E., George,R., Gonzalez,M., Guarin,H., Harris,N., Li,P.,
Liao,G., Miranda,A., Misra,S., Mungall,C.J., Nunoo,J., Pacleb,J.,
Paragas,V., Park,S., Phoumenavong,S., Wan,K., Yu,C., Lewis,S.E.,
Celisner,S. and Rubin,G.M.
BDGP/HMT AT Drosophila EST Project
Unpublished (2000)
Other ESTs: AT01136.5prime
Contact: Stapleton, M.
BDGP

TITLE JOURNAL
COMMENT Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: <http://www.fruitfly.org/EST>
hit genomic AE003746: arm:3R [19764572,19982668]
estimated-cyto:95D5-95F1: 04/07/2001
Plate: AT.11 row: C column: 12
High quality sequence stop: 631.
Location/Qualifiers
1..683
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="AT01136"
/clone_lib="AT Drosophila melanogaster adult testes POTB7"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
AT.121-AT.319: DH5-alpha Tona"
/note="Organ: ADULT testes; Vector: POTB7; Site:1. EORI:
Site:2: Xho1; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into POTB7. Plasmid cDNA library."

FEATURES
source

BASE COUNT 169 a 188 c 149 g 177 t
ORIGIN

alignment_scores:
Quality: 507.50 Length: 176
Ratio: 3.574 Gaps: 1
Percent Similarity: 80.682 Percent Identity: 57.955

alignment_block:
US-09-581-241-6 x BF492687/rev ..

Align seg 1/1 to reverse of: BF492687 from: 1 to: 683

```

373 lAlaLysValIleAspLeuAspThrLysThrLysThrLysProAsnArg 389
|||||
637 GCCAAGGTGATCGATCCGACACCGCAAGCTATTGCGGCCAACGAGC 588
|||||
389 gGlyGluValCysValLysGlyProMetLeuMetLysGlyTyrValAsp 406
|||||
587 CGCGAGCTTTGTTTAAAGCGACGCAATCATGAAGGCTACATCGAG 538
|||||

```



```

651 TGGCTCCACATTAAT 666
seq_name: gb_est2:BG414540
seq_documentation_block:
  LOCUS      BG414540      876 bp      mRNA      13-MAR-2001
  DEFINITION  HVSMEK0002N03f Hordeum vulgare testa/pericarp EST library
  HVCNNA0013 (normal) Hordeum vulgare cDNA clone HVSMEK0002N03f, mRNA
  sequence.
ACCESSION   BG414540
VERSION     BG414540.1 GI:13320091
KEYWORDS    EST.
SOURCE      barley.
ORGANISM    Hordeum vulgare
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
             ; Triticeae; Hordeum.
REFERENCE   1 (bases 1 to 876)
AUTHORS     Wing,R., Close,T.J., Kleinofe,A., Wise,R., Begum,D., Frisch,D., Yu
             ,Y., Anderson,H., Dale,J., Henry,D., Kernodle,S., Palmer,M., Rambo
             ,T., Saski,C., Schwartzbeck,J., Simmons,J., Choi,D.W., Main,D. and
             Wood,T.
TITLE       Development of a genetically and physically anchored EST resource
             for barley genomics
JOURNAL     Unpublished (2000)
COMMENT     Contact: Wing RA
             Clemson University
             100 Jordan Hall, Clemson, SC 29634, USA
             Tel: 864 656 7288
             Fax: 864 656 4293
             Email: rwing@clemson.edu
             Seq primer: AATTAACTTCACTCAAGG
             High quality sequence stop: 847.
FEATURES
  source     1..876
             /organism="Hordeum vulgare"
             /cultivar="Morex"
             /db_xref="taxon:4513"
             /clone="HVSMEK0002N03f"
             /clone_11b="Hordeum vulgare testa/pericarp EST library
             HVCNNA0013 (normal)"
             /tissue_type="testa/pericarp"
             /lab_host="TJUC121"
             /note="Vector: lambdaZAP. Site_1: EcoRI, Site_2: XhoI. For
             more details on library preparation and sequence analysis
             see http://www.genome.clemson.edu/projects/barley/ To
             order a clone see http://www.genome.clemson.edu/orders"
BASE COUNT  248 a 187 c 209 g 232 t
ORIGIN
alignment_scores:
  quality: 505.50      length: 215
  ratio: 3.082         gaps: 4
Percent Similarity: 76.279 Percent Identity: 49.767
alignment_block:
US-09-581-241-6 x BG414540 ..
Align seg 1/1 to: BG414540 from: 1 to: 876
333 LysGluIleGlyGluAlaValAlaArgArgPheAsnLeuProGlyValAr 339
      ::::::::::::::::::::::::::::::::::::
6 AGGAGCGCATGAGAGGTGGTGGCCAGAACTCCGCCAGCGGATCGT 55
339 gGlnGlyTyrGlyLeuThrGluThrThrSerAlaIleIleIleThrProG 356
      |||||||:::|||||:::|||||:::|||||:::
56 TCAGGCGCATGATGACTGACTGTGGG.....ATCATATCGCTTG 99
356 LuGlyAspAspLysPro.....GlyAlaSerGlyLysVal 367
      ||:::|||||:::|||||:::|||||:::
100 AATACCCAGAAAGAGACAGCGCTGTCATTTGGTCAACCGAACGCTT 149

```

```

368 ValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrLe 384
      |||:::|||||:::|||||:::|||||:::|||||:::
150 GTTGTGGAGTTGAGCAAAAGTTATCGAGTAGAAGCAAAACATCT 199
384 uGlyProAsnArgArgGlyGlyValLysValLysGlyPrometLeuMetL 401
      |:::|||||:::|||||:::|||||:::|||||:::
200 ACCGCCAAGTCAAGTACAGAAATCTGTATCCGAGGACACACATTAATGC 249
401 ysgGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGluG 417
      ::|||||:::|||||:::|||||:::|||||:::
250 AAGGGTATTTCACACATGTAGAGGCTACT...GAATTTACAAATCAGCA 296
418 GYTTPLeuHisThrGlyAspIleGlyTyrTyrAspGluGlyHisPh 434
      |||||||:::|||||:::|||||:::|||||:::
297 GATGTTTCATACCTGCTGATCTTGGATCTTGTAGAAAGAGCCAGCT 346
434 ePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 451
      ::|||:::|||||:::|||||:::|||||:::
347 TTTTTCGTCGATAGACTAAAGAGCTGATTAAGTACAAAGTTCCAGA 396
451 aLProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePhe 467
      ::|||:::|||||:::|||||:::|||||:::
397 TTGCACCTGCTGAGCTTGAAGATTAATCTTATCTCATCCAGAGATCTA 446
468 AspAlaGlyValAlaGlyValProAspProIleAlaGlyLysLeuProG 484
      |||||:::|||||:::|||||:::|||||:::
447 GACGCCGCTGTTATCCCGTCCCTGATGATGAAGCTGGGGAAGTTCTAT 496
484 yAlaValAlaValAlaLeuLysGlyLysSerMetThrGluLysGluValM 501
      |||:::|||||:::|||||:::|||||:::
497 TGGCTATGTGTGAGGTGACCTGACCTGATTAACCGAAGAGAGCTGCC 546
501 eCAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyGly 517
      :::::|||||:::|||||:::|||||:::
547 AAAAATTTATCGCCAAACAGGTCACTACTACAAAGAGCTGAGG...CGG 593
518 ValArgPheValAspGlyValProLysGlyLeuThrGlyLysIle 532
      |||:::|||||:::|||||:::|||||:::
594 GTTACTTGTGTGAGAGCTGCCAAATATCAGCCGACGCAAGATC 638
seq_name: gb_est1:AM691106
seq_documentation_block:
LOCUS      AM691106      656 bp      mRNA      15-JUN-2000
DEFINITION  NF041C08S11P1000 Development stem Medicago truncatula cDNA clone
NF041C08ST 5', mRNA sequence.
ACCESSION   AM691106
VERSION     AM691106.1 GI:7565842
KEYWORDS    EST.
SOURCE      barrel medic.
ORGANISM    Medicago truncatula
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
             Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliales;
             Medicago.
REFERENCE   1 (bases 1 to 656)
AUTHORS     He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
             ,C.-J., Flores,H.R., Imman,J.T., Weller,J.W., May,G.D. and Dixon
             ,R.A.
TITLE       Expressed Sequence Tags from the Samuel Roberts Noble Foundation
             Medicago truncatula stem library
JOURNAL     Unpublished (2000)
COMMENT     Contact: Dixon RA
             Plant Biology Division
             The Samuel Roberts Noble Foundation
             2510 Sam Noble Parkway, Ardmore, OK 73402, USA
             Tel: 580 221 7302
             Fax: 580 221 7380
             Email: radixon@noble.org
             Insert Length: 656 Std Error: 0.00
             Plate: 041 row: C column: 08
             Seq primer: TCACACAGAAACAGCTATGAC.

```


117 CCGGACAGTGTGGCTATGCTTGGCATTTGCTAAAGAACCTTCGAA 166
 359 AsplysProglialaSerGlyLysValaProleuphelysalalysVa 375
 167 ATTAATCTGGAGCATGTGGACTGTGTAGAAATGCTCGATGAATAT 216
 375 lileaspleusprhlylslsthrleuglyProasnarargrlygluv 392
 217 TGTGATCTGCTGATGACCAATCTCCCTAGAACCAATCTGAGAGA 266
 392 alycysValysglYprometleumetlysglyTyTyValaAspAsnProglu 408
 267 TTTGTATAGAGCGCATCAAAATTATGAAGGTTACCTAAAGATCCACAG 316
 409 AlathrArgluileleaspgluglylYrPrleuHsthrGlyAspII 425
 317 GCCACTAGCGGACAAATAGACAAAGAGGTTGATTACATAGCGGCGATAT 366
 425 eglYTyTyAspGlygluLysHisPhepheleValaAspArgleuLys 442
 367 TGTATATATGACATGATGATGAGCTTTTCATGTGTGATGCTTAAAG 416
 442 erleuilelYsTyTyglYTyTyValaProProalagluLys 458
 417 AATTGATAAATCAAGATTTCAAGTGGCTCTGCTGACCTGCAACCC 466
 459 ValleuLeuglnHisProasnllePheaspalaglyValaAlaglyValaPr 475
 467 CTCTTCCTCAATCATCCCAATTTTCAGATGCTGCTGTTGTTCAATGAA 516
 475 OASpProilealaglygluLeuProglYValaValaValleuLys 492
 517 AGACGACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 566
 492 lYlYSerMetThrGluLysgluValMetAspTyTyValaAlaSerGlnVal 508
 567 GCTCCCAATCTACTGATGAAGTCAAGATTCATCAACANAGAGGTTG 616
 seq_name: gb_est1:BE454671
 seq_documentation_block: 686 bp mRNA EST 02-MAR-2001
 LOCUS BE454671
 DEFINITION HVSMEN0094P16f Hordeum vulgare 5-45 DAP spike EST library
 HVCNDA0009 (5 to 45 DAP) Hordeum vulgare cDNA clone HVSMEN0094P16f,
 mRNA sequence.
 BE454671
 VERSION BE454671.2 GI:13189489
 KEYWORDS EST.
 SOURCE
 ORGANISM
 Hordeum vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
 ; Triticeae; Hordeum.
 1 (bases 1 to 686)
 Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu
 Y., Anderson,H., Dale,J., Henry,D., Kernodle,S., Palmer,M., Rambo
 T., Saski,C., Schwartzbeck,J., Simmons,J., Chol,D.W., Main,D. and
 Wood,T.
 TITLE
 Development of a genetically and physically anchored EST resource
 for barley genomics
 JOURNAL
 Unpublished (2000)
 COMMENT
 On Jul 26, 2000 this sequence version replaced gi:9463485.
 Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seg primer: AATTAACTCCCTCAATAAGG
 High quality sequence stop: 653.
 FEATURES
 Location/Qualifiers
 1..686

/organism="Hordeum vulgare"
 /cultivar="Morex"
 /db_xref="taxon:4513"
 /clone="HVSMEN0094P16f"
 /clone_lib="Hordeum vulgare 5-45 DAP spike EST library
 HVCNDA0009 (5 to 45 DAP)"
 /tissue_type="5-45 DAP Spike"
 /lab_host="SOLR"
 /note="Vector: lambdaZAP. Site 1: EcoRI; Site 2: XhoI. For
 more details on library preparation and sequence analysis
 see <http://www.genome.clemson.edu/projects/barley/> To
 order a clone see <http://www.genome.clemson.edu/orders>"
 BASE COUNT 156 a 212 c 203 g 115 t
 ORIGIN
 alignment_scores:
 Quality: 482.00 Length: 223
 Ratio: 2.957 Gaps: 2
 Percent Similarity: 73.094 Percent Identity: 43.049
 alignment_block:
 US-09-581-241-6 x BE454671 ..
 Align seg 1/1 to: BE454671 from: 1 to: 686
 314 lileaserglylYalProleuSerlysglylYleuglylYalValaI 330
 17 GTCATGTCCGGCGCGGCCCATGGGCAAGAGCTGACAGACGCTTCAT 66
 330 AARGAPheAsnLeuProglYValaArgGlyTyTyGlyLeuThrGluT 347
 67 GGCCAAGATCCCAACCGCGTCTGGGCGAGGAGGATGACGAGG 116
 347 hrThrSerAlaIlelelThrProgluLysAspLysPro..... 361
 117 CCGGACCGGTCTGCCATGCTGCTGCGCTTCCCAAGAGGCGCTTCAG 166
 362GlyAlaSerGlyLysValaProleuphelysalalysVa 375
 167 GTCAAGTCCGGGTGCTGCGGACGAGGTGTCGCAAGCGCGGCTCAAGAT 216
 375 lileaspleusprhlylslsthrleuglyProasnarargrlygluv 392
 217 CGTGCACCCCGACACCGCGGCTCTGCGCGGACGAGCGCGCAGA 266
 392 alycysValysglYprometleumetlysglyTyTyValaAspAsnProglu 408
 267 TGTGATCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAA 316
 409 AlathrArgluileleaspgluglylYrPrleuHsthrGlyAspII 425
 317 TCCACAGAACACCAATGACAGAGATGCTGCTGACACCGGACGACAT 366
 425 eglYTyTyAspGlygluLysHisPhepheleValaAspArgleuLys 442
 367 CGGCATCTTCACGACGACGAGATCTTCATTTGACAGGCTCAAG 416
 442 erleuilelYsTyTyglYTyTyValaProProalagluLys 458
 417 AGATCATCAAGTACAGGCTTCCAGGTGCGCGGACGAGCTGAGGCC 466
 459 ValleuLeuglnHisProasnllePheaspalaglyValaAlaglyValaPr 475
 467 CTCTTCATCACCACCGGAGATCAAGAGCGCGGCTGATGCGTGA 516
 475 OASpProilealaglygluLeuProglYValaValaValleuLys 492
 517 GGACCATCTTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 566
 492 lYlYSerMetThrGluLysgluValMetAspTyTyValaAlaSerGlnVal 508
 567 GCTTGAATATCCGAGATGATCAAAATTCGTGCGCAAGAGGTT 616

VERSION BF064865.2 GI:13266207
 KEYWORDS EST.
 SOURCE Hordeum vulgare
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
 1 (bases 1 to 806)
 REFERENCE 1 (bases 1 to 806)
 AUTHORS Wang, R., Close, T.J., Kleinbols, A., Wise, R., Begum, D., Frisch, D., Yu, Y., Anderson, H., Dale, J., Henry, D., Kernodle, S., Palmer, M., Rambo, T., Sasaki, C., Schwarzbeck, J., Simmons, J., Choi, D.W., Main, D. and Wood, T.
 TITLE Development of a genetically and physically anchored EST resource for barley genomics
 JOURNAL Unpublished (2000)
 COMMENT On Oct 17, 2000 this sequence version replaced gi:10841504.
 CONTACT: Wang RA
 CLEMSON UNIVERSITY Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seq primer: AATTACCTCCTACTAAAGG
 High quality sequence stop: 759.
 Location/Qualifiers
 1..806
 /organism="Hordeum vulgare"
 /cultivar="C116151 (M186)"
 /db_xref="taxon:4513"
 /clone_lib="HV_CBD0021D12f"
 /clone_lib="Hordeum vulgare seedling green leaf EST library HVCDA0005 (Erysiphe infected & control)"
 /tissue_type="seedling green leaf"
 /lab_host="SOLR"
 /note="Vector: lambdaZAP, Site_1: EcoRI, Site_2: XhoI. For more details on library preparation and sequence analysis see http://www.genome.clemson.edu/projects/barley/ To order a clone see http://www.genome.clemson.edu/orders"
 BASE COUNT 185 a 201 c 239 g 181 t
 ORIGIN
 alignment_scores:
 Quality: 479.00 Length: 211
 Ratio: 3.051 Gaps: 3
 Percent Similarity: 74.408 Percent Identity: 45.972
 alignment_block:
 US-09-581-241-6 x BF064865 ..
 Align seg 1/1 to: BF064865 from: 1 to: 806
 340 GlnGlyTyrGlyLeuThrGlnThrThSerAlaIleIleThrProG1 356
 |||||
 3 CAGGGGTACGGGCTGACCGAGACGGCCATCGCGCTCCACGACTC 52
 356 uG1AspAsp.....LysProGlyAlaSerGlyLysValValProLeuP 371
 :|||
 53 GGCCGAGGAGCGCCGCTACGGGACGGCGGCTCTGTGCGCCACA 102
 371 helysAlaValIleAspLeuAspThrLysLysThrLeuGlyProAsn 387
 :|||
 103 CCGAGGCGCAATGCTGACCGGAGACGAGCGCGCTGCGGTGAAC 152
 388 ArgArgGlyGlyValLysValLysGlyProMetLeuMetLysGlyTyrVa 404
 |||||
 153 CGCACCGGCGGCTGCTGATCAGGGGACCTACGTCATGAAGATACTT 202
 404 LAspAsnProGluAlaThrArgGluIleIleAspGluGlyTyrPLeuH 421
 |||||
 203 CAAGACACAGAGGACACAGTGCAGCGTGACGCCCAAGATGCTCA 252

421 IsThrGlyAspIleGlyTyrTyrAspGluGluLysHisPheIleVal 437
 |||||
 253 AGACCGGTGACCTATGCTACATAGACGAGATGGATCTCTTCGGGG 302
 438 AspArgLeuLysSerLeuIleLysTyrLysGlyTyrGluValProProAl 454
 |||||
 303 GACCGTGTGAAGAGTTGATCAAAATCAAAAGCTATCAGTGCCTCCAC 352
 454 acIleuGluSerValLeuLeuGlnHisProAsnIlePheAspAlaGly 471
 |||||
 353 AAGATTGGAGCTCTTCTGTACACCATCCAGAGGTTCCGATGCTG 402
 471 AlAlaGlyValProAspProIleAlaGlyLeuProGlyAlaValAl 487
 |||||
 403 TTATTCCTTCCACAGACGAGAGGTGTCAGTCCCATGCGCTATGTC 452
 488 ValLeuLysLysGlyLysSerMetThrGluLysGluValMetAspTyrVa 504
 |||||
 453 GTGAGAGAGAAAGAGGAGCAATTTGTCAGCCAGAGAGTATGAGATTGT 502
 504 lAlaSerGluValSerAsnAlaLysArgLeuArgGlyValArgPheV 521
 |||||
 503 GCGGAACACAGGTAGCACCCTTACACAGAGAGGTGAG...AAGTGGCGCTTG 549
 521 AlAspGluValProLysGlyLeuThrGlyLysIleAspLysAlaIle 537
 |||||
 550 TCACAGACATCCCCCAAGATGGCTGCGCAAGATATGTAGAGAA... 593
 538 ArgGluIleLeuLysLysProValAlaLysMet 548
 |||||
 594 ...GATCTTATCAAGCTCGCAGCTCCCAAGCTG 623
 seq_name: gb_est2:BB820931
 seq_documentation_block:
 LOCUS BB820931 772 bp mRNA EST 24-MAY-2001
 DEFINITION GM700013A20B2 Gm-r1070 Glycine max cDNA clone Gm-r1070-5044 3',
 mRNA sequence.
 ACCESSION BB820931
 VERSION BB820931.1 GI:10253165
 KEYWORDS EST.
 SOURCE soybean.
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 1 (bases 1 to 772)
 REFERENCE 1 (bases 1 to 772)
 AUTHORS Voklin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V., Erpelting, J., Raph, C., Shoop, E., Pardinas, J., Liu, L. and Lewin, H.
 TITLE A Functional Genomics Program for Soybean (NSF 9872565)
 JOURNAL Unpublished (1999)
 COMMENT Other-ESTs: AW458737 corresponding to Gm-cl016-6629 (5')
 Contact: Voklin, L.O., PI, A Functional Genomics Program for Soybean (NSF 9872565)
 Lewin, H. A., Director, Keck Center for Comparative and Functional Genomics
 University of Illinois
 Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
 Tel: (217) 244-6147
 Fax: (217) 333-4582
 Email: l.voklin@uiuc.edu
 This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134. For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genome.com
 Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'.
 Location/Qualifiers
 1..772
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="Gm-r1070-5044"

/clone.lib="Gm-r1070"
 /note="The library Gm-r1070 is a sequence-driven, rerecked set of 9,216 clones selected from cDNA libraries from various tissues and stages of development of soybean that represent 2,639 sequences from immature cotyledons, 1,770 from immature seed coats, 3,938 from flowers, and 869 from young pods. The 5' ESTs of the source clones from the different libraries was used to select singletons, or a representative of each contig, which were rerecked to form library Gm-r1070. The cDNA clones of the rerecked Gm-r1070 library were then sequenced at the 3' end. The contig analysis to select unique genes was performed by the Laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, http://www.cbc.umn.edu/research/projects/soybean/index.html. Rerecking was performed by Genome Systems, St. Louis, http://www.genomesystems.com, and 3 sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, http://www.life.uiuc.edu/biotech/keck.html. Note: The corresponding 5' EST from each clone in the Gm-r1070 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Genome Systems clone ID of the original cDNA library that is also listed under 'OTHER EST'."

BASE COUNT 195 a 190 c 143 g 214 t 30 others
 ORIGIN

alignment_scores:
 Quality: 477.50 Length: 243
 Ratio: 2.776 Gaps: 5
 Percent Similarity: 70.782 Percent Identity: 43.210

alignment_block:
 US-09-581-241-6 x BE820931/rev ..

Align seg 1/1 to reverse of: BE820931 from: 1 to: 772

```

307 AspLeuSerAsnLeuValGluIleAlaSerGlyValAlaProLeuSerly 323
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
747 GATNNNNCGTCGCGANNNGATCGTTGCCGCGAGAGAN 698
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
323 sgluilegylgualavalalargargpheaasleuproglyvalarg 340
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
697 ANNNNGATGAGGAGCATCAAGCTAGTCCCAAGTCCTCGTCAATGC 648
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
340 lnglytyrglyleuthrgluThrThrSerAlaIlelle..IleThrPro 355
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
647 AGGCAATCGGTTTAAGTACG.....TCTGCAGTTACCCGAAACCACTCCA 604
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
356 GluGlyspasplysproglyAlaSerGlyValValProLeuPhelely 372
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
603 GAGAGGCAAAATCAGTGGAGCAACAGTAACTGATACCAAAATATAGA 554
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
372 sAlaIleValIleAspLeuAspThrIlystThrLeuGlyProAsnArg 389
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
553 AGCCAAATTTGTGAAACCAAGGAAACGATGTTCCGCGTGAAC 504
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
389 rglygluValCysValIySGlyPrometleuMetlySGlyTyValAsp 405
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
503 AAGGAGAACTCTGGATCAGAGGACCTTATGTATGAAGGTTATTCGTG 454
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
406 AspProGluAlaThrArgGluIleIleAspGluGlyTyThrLeuHisTh 422
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
453 GACCAAAAGCACTTCAGCAACTTTG...GTGATGGGTGTTAAAGAC 407
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
422 rglAspIleGlyTyThrAspGluGlyHisPhePheIleValAsp 439
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
406 TGGGACCTCTGTATTTGTATACAAAGTTCTGTATGTTGAGATA 357
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
439 rglLeuIySerleuIleIySGlyTyrgluValProProAlaGlu 455
  
```

```

|||||.....|||||.....|||||.....|||||.....|||||.....
356 GGTGAAGAAGTTGATTAATACAAAGGCTACCAAGTGTGCTCCGCGAGAG 307
456 LeuGluSerValIleLeuGluHisProAsnIlePheAspAlaGlyValAl 472
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
306 CTAGAAAGTTACTCTATCTATCTATCTATCTATCTATCTATCTATCTAT 257
472 aglyValProAspProIleAlaGlyGluLeuProGlyValAlaValAl 489
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
256 TCCATACCTTCATAGTACGAGCCAGAGTCCATGCGCTTTGTGTGA 207
489 euIySGlySGlySerMetThrGluSGlyValMetAspTyValAla 505
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
206 GACAACCTCAAGATTCCTTGGCCAGACAGATTAATGATTGTTGCGC 157
506 SerIleValSerAsnAlaIySGlyValAlaArgPheValas 522
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
156 AAGAGGTTTCACCGTACAAAGAAATAAG...CGTATGACATTGTAA 110
522 pgluValProIySGlyLeuThrGlyIySGlyIleAspGlyValAlaIleArg 539
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
109 TTCCATACCCAGAAATGCTGCGGGAAGATP.....CTGAGA 72
539 luIleLeuIySGlyProValAlaIySGlyMet 548
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
71 AGGACTTAATAAAGTGGCGCTCTCTAGTTG 43
  
```

seq_name: gb_est1:AV558307

seq_documentation_block:

LOCUS AV558307 604 bp mRNA 07-SEP-2000
 DEFINITION AV558307 Arabidopsis thaliana green siliques Columbia Arabidopsis
 thaliana cDNA clone SQ094e05F 3', mRNA sequence.

ACCESSION AV558307
 VERSION AV558307.1 GI:8729733

KEYWORDS EST.
 SOURCE thale cress.

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 604)
 Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
 A large scale analysis of cDNA in Arabidopsis thaliana: Generation
 of 12,028 non-redundant expressed sequence tags from normalized and
 size-selected cDNA libraries

JOURNAL DNA Res. 7, 175-180 (2000)

COMMENT Contact: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizuekazusa.or.jp, URL: http://www.kazusa.or.jp/en/Plant/.

FEATURES
 source
 1..604
 location/Qualifiers
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="SQ094e05F"
 /clone_lib="Arabidopsis thaliana green siliques Columbia"
 /tissue_type="green siliques"
 /note="Vector: pBluescriptII SK-. Site_1: EcoRI, Site_2:
 XhoI"

BASE COUNT 165 a 129 c 148 g 162 t
 ORIGIN

alignment_scores:

Quality: 474.50 Length: 203
 Ratio: 3.142 Gaps: 2
 Percent Similarity: 74.384 Percent Identity: 47.291

alignment_block:

US-09-581-241-6 x AV558307 ..

Align seg 1/1 to: AV558307 from: 1 to: 604

```
304 AsplysTyraAspleSerAsnLeuValGluIleAlaSerGlyGlyAlaIpr 320
      :::::::::::::::::::: ::::::::::::::
3   AATTCCTATGATCTCTCCCTCGTATGATGCTTCTTCCGGTCAGCTCC 52
320 OlauSerLysGluIleGlyGluAlaValAlaArgArgPheAsnLeuProG 337
      :::::::::::::::::::: ::::::::::::::
53 TTTAGTTAAAGAACTTCAAGATAGTCTTCGCGCGTCTCCCTCAAGCCA 102
337 LyValArgGlnGlyTyrglyLeuThrGluThrThrSerAlaIleIle 353
      :::::::::::::::::::: ::::::::::::::
103 TCCTTGGCCAGGATATGTATGACGAGGAGCAGCAGTGTGTCAATG 152
354 Thr.....ProGluGlyAspAsplysProG1 362
      :::::::::::::::::::: ::::::::::::::
153 AGCCTTGCGTTCCGCTAAAGAACCGATACG.....ACAAATCAGG 193
362 YAlaSerGlyLysValValProLeuPheLysAlaLysValIleAspleuA 379
      :::::::::::::::::::: ::::::::::::::
194 TTCATGTGGGAGCTGTGCTCCGAAGCAGAGACTTAAGTGTTCACCTTG 243
379 sPrThrLysThrLeuGlyProAsnArgArgGlyGlyValLysValLys 395
      :::::::::::::::::::: ::::::::::::::
244 AGACACGCTCTCTCTGTGATACAAACCAACOTGAGATTGTATCCGC 293
396 GlyProMetLeuMetLysGlyTyrglyValAspAsnProGluAlaThrArg1 412
      :::::::::::::::::::: ::::::::::::::
294 GGTCAACAGATCATGAAGAAAGTACTGACGATCCAGAACGCCACTTCAGC 343
412 uIleIleAspGluGluGlyTyrrPheuHISThrGlyAspIleGlyTyrrA 429
      :::::::::::::::::::: ::::::::::::::
344 AACATAGACCAAGAAAGTGTGCTTCACACAGAGACATTGATATGTG 393
429 sPgIuGluLysHISPhePheIleValAspArgLeuLysSerLeuIleLys 445
      :::::::::::::::::::: ::::::::::::::
394 ATGAAGATGATGATGATTTTCATTTGTGATGACTTAAGAAGTCATCAG 443
446 TyrrLysGlyTyrglyValProProAlaGluLeuGluSerValLeuLeuG1 462
      :::::::::::::::::::: ::::::::::::::
444 TTCAAAGGCTTTCAGTCCCTCCAGCTGAGCTAGAGATTGCTAATCAA 493
462 nHISProAsnIlePheAspAlaGlyValAlaGlyValProAspProIleA 479
      :::::::::::::::::::: ::::::::::::::
494 TCACCATTTCAATTGCCATGACGCTGTGTTCCGCAAAATGATGAGTGC 543
479 IaGlyGluLeuProGlyAlaValValValLeuLysGlyLysSerMet 495
      :::::::::::::::::::: ::::::::::::::
544 CTGGGGAAGTTCGCGTGTTCGTGCGATCAATGGAATGATATC 593
496 ThrGluLys 498
      ::::::::::::::
594 ACTGAAGAA 602
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2001, 01:53:50 ; Search time 25.81 Seconds

(without alignments)
778.471 Million cell updates/sec

Title: US-09-581-241-4

Perfect score: 2823
Sequence: 1 MEMMENDEIVYGPPEFYPI.....TGKIDSKAIRLEIKKPVAKM 548

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2814	99.7	548	1	LUCI_LUCIA
2	2684	95.1	548	1	LUCI_LUCR
3	1945.5	68.9	550	1	LUCI_PHOY
4	768.5	27.2	544	1	ACLI_PENCR
5	760.5	26.9	544	1	ACLI_PENCR
6	758.5	26.9	547	1	ACLI_PENCR
7	748.5	26.5	542	1	ACLI_TORAC
8	747	26.5	553	1	ACLI_TORAC
9	734.5	26.0	543	1	ACLI_TORAC
10	733	25.0	543	1	ACLI_TORAC
11	732	25.0	561	1	ACLI_TORAC
12	727.5	25.8	545	1	ACLI_TORAC
13	726	25.7	562	1	ACLI_TORAC
14	720.5	25.5	561	1	ACLI_TORAC
15	712	25.2	536	1	ACLI_TORAC
16	689.5	24.4	569	1	ACLI_TORAC
17	674	23.9	537	1	ACLI_TORAC
18	587.5	20.8	293	1	ACLI_TORAC
19	572.5	20.3	560	1	ACLI_TORAC
20	569	20.2	561	1	ACLI_TORAC
21	481	17.0	562	1	ACLI_TORAC
22	463	16.4	543	1	ACLI_TORAC
23	461	16.3	572	1	ACLI_TORAC
24	444	15.7	548	1	ACLI_TORAC
25	427.5	15.1	522	1	ACLI_TORAC
26	368	13.0	698	1	ACLI_TORAC
27	366	13.0	683	1	ACLI_TORAC
28	359	12.7	699	1	ACLI_TORAC
29	357	12.6	492	1	ACLI_TORAC
30	351.5	12.5	699	1	ACLI_TORAC
31	350	12.4	699	1	ACLI_TORAC
32	346.5	12.3	683	1	ACLI_TORAC
33	338	12.0	698	1	ACLI_TORAC

34	334.5	11.8	539	1	DBE_BACSU
35	325	11.5	711	1	ICPD_HUMAN
36	324.5	11.5	670	1	ICPD_RAT
37	323.5	11.5	6359	1	BACC_BACLI
38	321	11.4	697	1	ICPF_HUMAN
39	317.5	11.2	694	1	ICPF_YEAST
40	312	11.1	697	1	ICPF_RAT
41	311.5	11.0	543	1	VIBE_VIBCH
42	308.5	10.9	711	1	ICPD_MOUSE
43	308.5	10.9	744	1	ICPD_MOUSE
44	306	10.8	5255	1	BACA_BACLI
45	305.5	10.8	700	1	ICP1_YEAST

ALIGNMENTS

RESULT 1
LUCI_LUCIA STANDARD: PRT: 548 AA.
AC 001158:
DT 01-APR-1993 (rel. 25, Created)
DT 01-APR-1993 (rel. 25, Last sequence update)
DT 20-AUG-2001 (rel. 40, Last annotation update)
DE LUCIFERIN 4-MONOOXYGENASE (EC 1.13.12.7) (LUCIFERASE).
OS Luciola lateralis (Firefly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga; Elateriformia; Cantharoidae; Lampyridae; Luciola.
OX NCBI_TaxID=7052;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92305054; PubMed-1610896;
RA Tatsumi H., Kajiyama N., Nakano E.;
RT "Molecular cloning and expression in Escherichia coli of a cDNA clone encoding luciferase of a firefly, Luciola lateralis.";
RL Biochim. Biophys. Acta 1131:161-165(1992).
CC - FUNCTION: PRODUCES GREEN LIGHT.
CC - CATALYTIC ACTIVITY: LUCIFERIN + O(2) + ATP = OXIDIZED LUCIFERIN + CO(2) + H(2)O + AMP + PYROPHOSPHATE + LIGHT.
CC - CORFACTOR: REQUIRES MAGNESIUM.
CC - SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
CC - SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

EMBL: X66919; CAA47358.1; -
PIR: S23437; S23437.
DR HSSP: P08659; LUCI.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00342; MICROBODIES-CTER; 1.
DR PROSITE: PS00455; AMP-BINDING; 1.
KW Oxidoreductase; Monooxygenase; Photoprotein; Luminescence; Magnesium; Peroxisome.
FT SITE 546 548 MICROBODIES TARGETING SIGNAL (POTENTIAL).
SQ SEQUENCE 548 AA: 60125 MW; ACQ2P9320BB6DA6 CRC64;

Query Match 99.7%; Score 2814; DB 1; Length 548;
Best Local Similarity 99.6%; Pred. No. 1.4e-188;
Matches 546; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEMMENDEIVYGPPEFYPIEGSGAGOLRKTYMDRYAKGATATNALTGVDYTAETYE 60
DB 1 MEMMENDEIVYGPPEFYPIEGSGAGOLRKTYMDRYAKGATATNALTGVDYTAETYE 60

```

OY 61 KSCCLGALKNKGVLVNDGRIALCSENCEEEFPIVLAGLFGVGAAPNEITTLRELHSL 120
DB 61 KSCCLGALKNKGVLVNDGRIALCSENCEEEFPIVLAGLFGVGAAPNEITTLRELHSL 120
OY 121 GISKPTIVFSSKGLDKVITVOKTVAIKTIVILDSKVYRGYOSMNFIKKNTPOGFKG 180
DB 121 GISKPTIVFSSKGLDKVITVOKTVAIKTIVILDSKVYRGYOSMNFIKKNTPOGFKG 180
OY 181 SSRKTIVANRKEOYALLMNSSGSTGLPKGVQVLTHTENLVTRSHARDPIYGNVSPGTAIL 240
DB 181 SSRKTIVANRKEOYALLMNSSGSTGLPKGVQVLTHTENLVTRSHARDPIYGNVSPGTAIL 240
OY 241 TVPFEHGFEGMTLLGYLTCGFRIVMLTKPEDEFTLTDYKSSVILVPTLPAIILNRS 300
DB 241 TVPFEHGFEGMTLLGYLTCGFRIVMLTKPEDEFTLTDYKSSVILVPTLPAIILNRS 300
OY 301 ELLDKYDLSNLVEIASGAPLSKEIGAVARFNLPGVRGQYGLTETTSALITPEGDDK 360
DB 301 ELLDKYDLSNLVEIASGAPLSKEIGAVARFNLPGVRGQYGLTETTSALITPEGDDK 360
OY 361 PGASGKVVPLFKAVIDLDTRKKTIGPNRGEVGVKGMKMGVYDNPATREITIDEGMWL 420
DB 361 PGASGKVVPLFKAVIDLDTRKKTIGPNRGEVGVKGMKMGVYDNPATREITIDEGMWL 420
OY 421 HTGDIGYDEKHFIVDRLSLIKYGVOVPALESVLLQHPNIPDAGVAGVDPPIAG 480
DB 421 HTGDIGYDEKHFIVDRLSLIKYGVOVPALESVLLQHPNIPDAGVAGVDPPIAG 480
OY 481 ELPGAVVVLKKGSMTEKEVMDYVASOVSNARKRLRGVRFVDEVPKGLTGKIDKAIKEI 540
DB 481 ELPGAVVVLKKGSMTEKEVMDYVASOVSNARKRLRGVRFVDEVPKGLTGKIDKAIKEI 540
OY 541 LKRPVAKM 548
DB 541 LKRPVAKM 548

RESULT 2
LUC1_LUCR STANDARD; PRT; 548 AA.
AC P13129;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE LUCIFERIN 4-MONOOXYGENASE (EC 1.13.12.7) (LUCIFERASE).
OS Luciola cruciata (Japanese firefly) (Genji firefly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Elateriformia; Cantharidea; Lampyridae; Luciola.
OX NCBI_TaxID=7051;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89326143; PubMed=2473944;
RA Masuda T., Tatsumi H., Nakano E.;
RT "Cloning and sequence analysis of cDNA for luciferase of a Japanese
  firefly, Luciola cruciata.";
RL Gene 77:265-270(1989).
CC -|- FUNCTION: PRODUCES GREEN LIGHT WITH A WAVELENGTH OF 544 NM.
CC -|- CATABOLIC ACTIVITY: LUCIFERIN + O(2) + ATP -> OXIDIZED LUCIFERIN +
  CO(2) + H(2)O + AMP + PYROPHOSPHATE + LIGHT.
CC -|- COFACTOR: REQUIRES MAGNESIUM.
CC -|- SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
  FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).

```

```

CC EMBL: M26194; AAA29135.1;
DR PIR: J50181; J50181.
DR HSSP: P08659; IIC1.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00342; MICROBODIES_CTER; 1.
DR PROSITE: PS00455; AMP-BINDING; 1.
KW Oxidoreductase; Monooxygenase; Photoprotein; Luminescence; Magnesium;
FT SITE 546 548 MICROBODY TARGETING SIGNAL (POTENTIAL).
SQ SEQUENCE 548 AA; 60017 MW; 2052D6189E79109F CRC64;

Query Match          95.1%; Score 2684; DB 1; Length 548;
Best Local Similarity 93.4%; Pred. No. 1.6e-179;
Matches 512; Conservative 26; Mismatches 10; Indels 0; Gaps 0;

OY 1 MENNENDENTIVYGPPEYPIEESGAGALRKRYMDRVAKLGAIAFTNALGVDTYAEYLE 60
DB 1 MENNENDENTIVYGPKPPIEESGAGTQLRKRYMERAKLGAIAFTNAVGTGVDSYAEYLE 60
OY 61 KSCCLGALKNKGVLVNDGRIALCSENCEEEFPIVLAGLFGVGAAPNEITTLRELHSL 120
DB 61 KSCCLGALKNKGVLVNDGRIALCSENCEEEFPIVLAGLFGVGAAPNEITTLRELHSL 120
OY 121 GISKPTIVFSSKGLDKVITVOKTVAIKTIVILDSKVYRGYOSMNFIKKNTPOGFKG 180
DB 121 GISKPTIVFSSKGLDKVITVOKTVAIKTIVILDSKVYRGYOSMNFIKKNTPOGFKG 180
OY 181 SSRKTIVANRKEOYALLMNSSGSTGLPKGVQVLTHTENLVTRSHARDPIYGNVSPGTAIL 240
DB 181 SSRKTIVANRKEOYALLMNSSGSTGLPKGVQVLTHTENLVTRSHARDPIYGNVSPGTAIL 240
OY 241 TVPFEHGFEGMTLLGYLTCGFRIVMLTKPEDEFTLTDYKSSVILVPTLPAIILNRS 300
DB 241 TVPFEHGFEGMTLLGYLTCGFRIVMLTKPEDEFTLTDYKSSVILVPTLPAIILNRS 300
OY 301 ELLDKYDLSNLVEIASGAPLSKEIGAVARFNLPGVRGQYGLTETTSALITPEGDDK 360
DB 301 ELLDKYDLSNLVEIASGAPLSKEIGAVARFNLPGVRGQYGLTETTSALITPEGDDK 360
OY 361 PGASGKVVPLFKAVIDLDTRKKTIGPNRGEVGVKGMKMGVYDNPATREITIDEGMWL 420
DB 361 PGASGKVVPLFKAVIDLDTRKKTIGPNRGEVGVKGMKMGVYDNPATREITIDEGMWL 420
OY 421 HTGDIGYDEKHFIVDRLSLIKYGVOVPALESVLLQHPNIPDAGVAGVDPPIAG 480
DB 421 HTGDIGYDEKHFIVDRLSLIKYGVOVPALESVLLQHPNIPDAGVAGVDPPIAG 480
OY 481 ELPGAVVVLKKGSMTEKEVMDYVASOVSNARKRLRGVRFVDEVPKGLTGKIDKAIKEI 540
DB 481 ELPGAVVVLKKGSMTEKEVMDYVASOVSNARKRLRGVRFVDEVPKGLTGKIDKAIKEI 540
OY 541 LKRPVAKM 548
DB 541 LKRPVAKM 548

RESULT 3
LUC1_PROPY STANDARD; PRT; 550 AA.
AC P08659;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE LUCIFERIN 4-MONOOXYGENASE (EC 1.13.12.7) (LUCIFERASE).
OS Photinus pyralis (North American firefly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Elateriformia; Cantharidea; Lampyridae; Photinus.
OX NCBI_TaxID=7054;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=87144243; Pubmed=3821727;
RA de Wet J.R., Wood K.V., Deluca M., Helinski D.R., Subramani S.;
RT "Firefly luciferase gene: structure and expression in mammalian
RL cells.";
RL Mol. Cell. Biol. 7:725-737(1987).
RN
[2]
RP SUBCELLULAR LOCATION.
RX MEDLINE=87204117; Pubmed=3554235;
RA Keller G.-A., Gould S., de Luca M., Subramani S.;
RT "Firefly luciferase is targeted to peroxisomes in mammalian cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:3264-3268(1987).
RN
[3]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=96398615; Pubmed=8805533;
RA Conti E., Franks N.P., Brick P.;
RT "Crystal structure of firefly luciferase throws light on a
RL superfamily of adenylate-forming enzymes.";
RL Structure 4:287-298(1996).
RN
[4]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX MEDLINE=99007339; Pubmed=9788915;
RA Franks N.P., Jenkins A., Conti E., Lieb W.R., Brick P.;
RT "Structural basis for the inhibition of firefly luciferase by a
RL general anesthetic.";
RL Biophys. J. 75:2205-2211(1998).
CC -1- FUNCTION: PRODUCES GREEN LIGHT WITH A WAVELENGTH OF 562 NM.
CC -1- CATALYTIC ACTIVITY: LUCIFERIN + O(2) + ATP = OXIDIZED LUCIFERIN +
CC CO(2) + H(2O) + AMP + PYROPHOSPHATE + LIGHT.
CC -1- COFACTOR: REQUIRES MAGNESIUM.
CC -1- SUBCELLULAR LOCATION: PEROXISOMAL.
CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M15077; AAA29795.1; -
DR EMBL; X84848; CAA59283.1; -
DR EMBL; U03687; AAA03561.1; -
DR EMBL; U89934; AAB64396.1; -
DR EMBL; U89935; AAB64399.1; -
DR PIR; A26772; A26772.
DR PDB; 1LC1; 26-MAR-97.
DR PDB; 1BA3; 11-NOV-98.
DR InterPro; IPR000873; AMP-Bind.
DR Pfam; PF00501; AMP-binding.1.
DR PROSITE; PS00342; MICROBODIES_CTRF; 1.
DR PROSITE; PS00455; AMP-BINDING; 1.
DR Oxidoreductase; Monooxygenase; Photoprotein; Luminescence; Magnesium;
DR Peroxisome; 3D-structure.
DR SITE 548 550 MICROBODIE TARGETING SIGNAL.
DR SEQUENCE 550 AA; 60745 MW; E380FCE9D56ACDCE CMC6;
SO

```

```

DB 121 SQPTVFWFSKGLQKILANVQKLPILQKIIIMDSKTDYGFQSMYTFVSHLPFGFNEXD 180
OY 163 FETVEVNRKEQVALIMNNSGSGTGLKRGVOLTENLVTRSHARDPIYGNQVSPGAILTV 242
DB 181 FPPESFDRDKTIALIMNSSGSTGLPRGVALPHRTACVRSRARDPIFGNQLIPDIALISV 240
OY 243 VPFHNGFCMFTTLGLTGCFRIVMLTKPDEFELTADYKCSSVILVPTLPAILNRSEL 302
DB 241 VPFHNGFCMFTTLGLTGCFRIVMLTKPDEFELTADYKCSSVILVPTLPAILNRSEL 300
OY 303 IDKYDLSNLHEIASGAPLSEKGEAVARFNLPGVROGYGLTETTSAILIIPEDGDKRG 362
DB 301 IDKYDLSNLHEIASGAPLSEKGEAVARFNLPGVROGYGLTETTSAILIIPEDGDKRG 360
OY 363 ASGKVVPLFKAVIIDLTKTGTGPNRRCGVCKGPMKMGVYNDPEAREITDESGMLHT 422
DB 361 AVGVVPEFEAKVVDLDGKLTGVNQRELGVGRPMKMGVYNNPEATNALIDKDGMLHS 420
OY 423 GDIGYYDEEKHFPIYDRLSLTKYGYOVPPAELSVLLQHPNIFDAGVAGVDPPIAGEL 482
DB 421 GDIAWDEDEHFFIYDRLSLTKYGYOVPPAELSVLLQHPNIFDAGVAGVDPPIAGEL 480
OY 483 PGAVVVLKKGKSMTEKEVADYVAVSOVSNARKLRGVRFEVDEPKGTGKIDKAIKREILK 542
DB 481 PAAVVVLEHGKTMTEKEIYDVAVSOVTTAKKLRGVVFVDEVPKGTGKIDKAIKREILI 540
OY 543 K 543
DB 541 K 541

```

RESULT 4

ID	ACLN_PETCR	STANDARD:	PRT:	544 AA.
AC	P14912:			
DT	01-APR-1990 (Rel. 14, Created)			
DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	4-COMMARATE-COA LIGASE 1 (EC 6.2.1.12) (4CL 1) (4-COMMAROYL-COA SYNTHASE 1).			
GN	4CL OR 4CL-1.			
OS	Petroselinum crispum (Parsley) (Petroselinum hortense).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
OC	Asteridae; euasterids II; Apiales; Apiaceae; Petroselinum.			
OX	NCBI_TaxID=4043;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89005119; Pubmed=3169018;			
RA	Lozoya E., Hoffmann H., Douglas C., Schulz W., Scheel D.,			
RA	Hahlbrock K.;			
RT	"Primary structures and catalytic properties of isoenzymes encoded by			
RT	the two 4-commarate:COA ligase genes in parsley.";			
RL	Eur. J. Biochem. 176:661-667(1988).			
RN	[2]			
RP	SEQUENCE OF 1-8 FROM N.A.			
RA	Douglas C., Hoffmann H., Schulz W., Hahlbrock K.;			
RT	"Structure and elicitor or U.V.-light-stimulated expression of two			
RT	4-commarate:COA ligase genes in parsley.";			
RL	EMBO J. 6:1189-1195(1987).			
CC	-1- CATALYTIC ACTIVITY: ATP + 4-COMMARATE + COA = AMP + DIPHOSPHATE +			
CC	4-COMMAROYL-COA.			
CC	-1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOIC			
CC	METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.			
CC	-1- INDUCTION: TRANSCRIPTIONALLY ACTIVATED BY BOTH FUNGAL ELICITOR AND			
CC	ULTRAVIOLET IRRADIATION.			
CC	-1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME			
CC	FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			

QY 442 SLIRKGVPPAPLESVLLQHPNIPDAGVAGVDPPIAGELPGAVVYLKKKSKTEKEVM 501
 DB 439 EIRYKGVAPALELLTHPTISDAVPMDEKAGVPAVAVRTNGFTTBEERK 498
 QY 502 DYVASOVNNAKRLRGVFEVDEVPKGLTKIDGKAIR 538
 DB 499 QVSKQVFFYKRI-FRVFVDAIRKSPSGKILRRDLR 534

RESULT 6
 ID 4CL1_TOBAC STANDARD; PRT; 547 AA.
 AC 024145;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE 4-COMARATE-COA LIGASE 1 (EC 6.2.1.12) (4CL 1) (4-COMAROYL-COA SYNTHASE 1).
 GN 4CL1.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96416441; PubMed=8819324;
 RA Lee D., Douglas C.J.;
 RT "Two divergent members of a tobacco 4-coumarate:coenzyme A ligase (4CL) gene family. cDNA structure, gene inheritance and expression, and properties of recombinant proteins.";
 RL Plant Physiol. 112:193-205(1996).
 CC -1- CATALYTIC ACTIVITY: ATP + 4-COMARATE + COA = AMP + DIPHOSPHATE + 4-COMAROYL-COA.
 CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U50845; AAB18637.1; -;
 DR HSSP: P08659; ILCI.
 DR Mendel: 24477; Nicita:1179;24477.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 KW Ligase; Phenylpropanoid metabolism; Multigene family.
 SQ SEQUENCE 547 AA; 59842 MW; 154DF6D684E3F51B CRC64;

Query Match 26.9%; Score 758.5; DB 1; Length 547;
 Best Local Similarity 35.1%; Pred. No. 1.9e-45;
 Matches 189; Conservative 89; Mismatches 205; Indels 55; Gaps 12;

QY 32 YMDRAKAGALAFN-----ALTGVD--YTYAEYLRKSC-CIGELAKNYGLVVDG 78
 DB 25 YIPKHLPHSHCFENISFSSRPCLINGANDOIITYAE-VELTCKKAVGLNKIGIQOKD 83
 QY 79 RALSCENCEEFFIPLVLAGLGVAPATNEIYTLRELVHSLGSLPTIVSSKGLDKV 138
 DB 84 TIMILLPSPEVFAFMGASVYGAISTMANPLFTPAEYVQAKAKSSAIIITQSCFVGKV 143
 QY 139 ITVQKTVAKITIVILDSKVDYRGYSMDNFKTKNTPQG-----FGSSFKYEVN-R 190
 DB 144 -----KDYASENDVYKVICIDSAPEGCLHFSLTLQSDHEIEIPEVXIQ 184

QY 191 KEQVALIMNSGSTGPKGVOLTHENLYTRFSHARDPIYGN-QVSPGATLTVPRHHGF 249
 DB 185 PDVVALLPVSSGTGTGLPKGMVTHKGLTVSAQOVDEGANLIYMSBDVLCVLFH-- 242
 QY 250 GMFTTGLYTCGFR-----IVMLTKFEDEFELTKLDYKSSVILVPTLPAIINRSELDK 305
 DB 243 -IYSLNSILLGLRGVGAAILIMQKFDIAFLEIQLQKYSIQEPVPYLAIAKSPYDS 301
 QY 306 YDLSNLVETASGAPLSKEIGEAVARFNLPGVROGYGLTETTSAILI-----TEBGDK 360
 DB 302 YDLSVRTVWSGAPLQKELEDAVRFKFPNAKIGQGYGTEAGPVLAMLAKRPFEDK 361
 QY 361 PGASGVVPLPFAKVIDLDTTKTLGPNRGEQVCKGPMYAKGVONPEATREIDEQWL 420
 DB 362 SGACGTVNRNAEMKIVDPDTCGSLPRNQPELCTIKGDOIIMKGLNDPEATRTTIDKGM 421
 QY 421 HTGDIGYDEEKEHFTVDRLSLIRKGYVPPAELESVLLQHPNIPDAGVAGVDPPIAG 480
 DB 422 HTGDIGFIDEDELFTVDRKLKELIKYKGFQVAPAEIEALLNHPNISDAVAVPMKDEQAG 481
 QY 481 ELPGAIVYLKKKSKMTEKEVMYASQVSNARKLRGCVFVDEVPKGLTKIDGKAIR 538
 DB 482 EVPAEVVRNSGSAITDEVKDFISKQVIFYKRVK-RVFEVETVPKSPSGKILRRDLR 538

RESULT 7
 ID 4CL2_TOBAC STANDARD; PRT; 542 AA.
 AC 024146;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE 4-COMARATE-COA LIGASE 2 (EC 6.2.1.12) (4CL 2) (4-COMAROYL-COA SYNTHASE 2).
 GN 4CL2.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96416441; PubMed=8819324;
 RA Lee D., Douglas C.J.;
 RT "Two divergent members of a tobacco 4-coumarate:coenzyme A ligase (4CL) gene family. cDNA structure, gene inheritance and expression, and properties of recombinant proteins.";
 RL Plant Physiol. 112:193-205(1996).
 CC -1- CATALYTIC ACTIVITY: ATP + 4-COMARATE + COA = AMP + DIPHOSPHATE + 4-COMAROYL-COA.
 CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U50846; AAB18638.1; -;
 DR HSSP: P08659; IBA3.
 DR Mendel: 24478; Nicita:1179;24478.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 KW Ligase; Phenylpropanoid metabolism; Multigene family.
 SQ SEQUENCE 542 AA; 59479 MW; CB5579A9BDFC003 CRC64;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae: euasterids I; Solanales; Solanaceae; Solanum.
 OC NCBI_TaxID=4113;
 RN [1]
 RX MEDLINE=91217100; PubMed=2022667;
 RA Becker-Andre M., Schulze-Lefert P., Hahlbrock K.;
 RT "Structural comparison, modes of expression, and putative cis-acting
 elements of the two 4-coumarate: COA ligase genes in potato.";
 RL J. Biol. Chem. 266:8551-8559(1991).
 CC -1- CATALYTIC ACTIVITY: ATP + 4-COUMARATE + COA = AMP + DIPHOSPHATE +
 CC 4-COUMAROYL-COA.
 CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
 CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M62755; AAA33842.1; -.
 DR PIR: A39827; A39827.
 DR HSSP: P08659; ILCT.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 DR Ligase; Phenylpropanoid metabolism; Multigene family.
 SO SEQUENCE 545 AA; 59619 MW; D6183683B774BA71 CRC64;

Query Match 26.08; Score 734.5; DB 1; Length 545;
 Best Local Similarity 34.28; Pred. No. 9e-44;
 Matches 185; Conservative 93; Mismatches 202; Indels 61; Gaps 12;

QY 32 YMDRYAKGAIATFN-----ATGVD--YTVAEYLEKSCCLGALKNGVAVDGR 79
 DB 23 YTPKPLPHSYCFENLSEFNSRPLIDANDRITYTVAEVELTSKRVAVGLKKLIDQKDT 82
 QY 80 IALSCENCEEFPIVLAGLFTGVAVAPTNEIYTLRELVHSIGISKPITVSSKGLDKVI 139
 DB 83 IMILLPNEEVPVAFIGASYIGALISTMANPLETPAVYKQAKASAKIVITQACFACKV- 141
 QY 140 TVQKTVTAIKTIVILDSKVDRYQSDMNFK-----KNTPOG-----FKGSSFKTVEV 188
 DB 142 -----KDY-----AIENDLKVICVDSVPGCYHCFSELIQSDHEIPDV 179
 QY 189 N-RKQVALINSSGSGTGLPRGYQTLHNTLRFSSHADPIYGN-QVSPGAILITVYFPH 246
 DB 180 KIQPDVVALPYSSGTGLPRGYMLTHKGLTVYAQOVGENANLIMMSDDVLCVLPLE 239
 QY 247 HGFMTLLGLYTCGFR-----IVMLTKFEDETFKTLQDYKCSSVITVPLFALINSEL 302
 DB 240 H---IYLSNVSLCALRGAILLMOKFDIAQLELLPKHVTITGPRPPIVLAIASPL 296
 QY 303 LDKYLSMLVEIASGGLPLSKLKEIGEAVARRNPGVROGYLTETTSATIT-----TPEG 357
 DB 297 VDNYSLSVRYVMSGAFLGKLEDAVARAKFPNKLGGYGMTEAGPLAMCLAFAPKPE 356
 QY 358 DDKPASKVYVPLFAKYIDIDTKKTGPNRRGCVGYGPMIMGYVNPATREIIDE 417
 DB 357 DIKSGACGTIVANAMKATVDPDGTGSLPRNPGELCTINGDQIMKGYLMDPATIATTEKE 416
 QY 418 GMLATGDIYDYDEKHFIVRLKSLIKYKGOYPPAELESVLLQHPNIPDAGVAGYDP 477
 DB 417 GMLHGDIGDIDDELFLVRLKELIKYKGOYPALEALLINHPDISDAVAVPMIDE 476
 QY 478 TAGELPGAVVYLKKKSKTEKEVMDYVAVSOVSMKRLGVRFEVDEVKGLTKIDGKAI 537
 DB 537

DB 477 OAGEVPVAFVNSNGSTTEDEVKDFISKQVIFYKRIK-RVFEVETVPKSPGKILRDL 535
 QY 538 R 538
 DB 536 R 536
 DB 536 R 536
 RESULT 10
 ID 4CL3.ARATH STANDARD; PRT; 561 AA.
 AC 09S777;
 DT 20-AUG-2001 (Rel. 40; Created)
 DT 20-AUG-2001 (Rel. 40; Last sequence update)
 DT 20-AUG-2001 (Rel. 40; Last annotation update)
 DE 4-COUMARATE--COA LIGASE 3 (EC 6.2.1.12) (4CL 3) (4-COUMAROYL-COA
 DE SYNTHASE 3)
 GN 4CL3 OR AT16G5060 OR F16G16.6.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=99348176; PubMed=10417722;
 RA Erling J., Buettner D., Wang Q., Douglas C.J., Somsich I.E.,
 RA Kombrink E.;
 RT "Three 4-coumarate:coenzyme A ligases in Arabidopsis thaliana
 RT represent two evolutionarily divergent classes in angiosperms.";
 RL Plant J. 19:9-20(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.D., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Cressy T.H., Dewar K.,
 RA Dunn P., Elgu P., Feldhlyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremetska I., Kutz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,
 RA Maltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Tortum M.J., Town C.D.,
 RA Uteback T., Van Aken S., Vayberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 408:816-820(2000).
 CC -1- CATALYTIC ACTIVITY: ATP + 4-COUMARATE + COA = AMP + DIPHOSPHATE +
 CC 4-COUMAROYL-COA.
 CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
 CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF106087; AAD47194.1; -.
 DR EMBL: AF106088; AAD47195.1; -.
 DR EMBL: AC009360; AAF06039.1; -.
 DR HSSP: P08659; ILCT.

Db 473 LEALLIGHPDITDVAVAMKEEAGEVPVAFVVKSKDSELSDDQFVSKQVFFYKRI- 531
 QY 516 GGVRFVDEVPKGLTGKIDGKAIR 538
 Db 532 NKVFETESIPKAPSGKILRKDLR 554

RESULT 15

4CL2_ARATH STANDARD; PRT; 556 AA.
 ID 4CL2_ARATH 095725: 091U35:
 AC 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE 4-COMARATE-COA LIGASE 2 (EC 6.2.1.12) (4CL 2) (4-COMARoyL-COA
 DE SYNTHASE 2).
 GN 4CL2 OR AT3G21240 OR MXL8.10.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
 OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae;
 OC eustosids II: Brassicales: Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV_COLUMBIA;
 RX MEDLINE=9348176; PubMed=10417722;
 RA Ehtling J., Buetelner D., Wang Q., Douglas C.J., Somasich I.E.,
 RA Kombrink E.;
 RT "Three 4-coumarate:coenzyme A ligases in Arabidopsis thaliana
 RT represent two evolutionarily divergent classes in angiosperms.";
 RL Plant J. 19:9-20(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV_COLUMBIA;
 RX MEDLINE=20277480; PubMed=10819329;
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
 RT features of the regions of 4,504,864 bp covered by sixty P1 and YAC
 RT clones.";
 RL DNA Res. 7:131-135(2000).
 CC -1- CATALYTIC ACTIVITY: ATP + 4-COMARATE + COA = AMP + DIPHOSPHATE +
 CC 4-COMAROYL-COA.
 CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
 CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF106085; AAD47192.1; -
 DR EMBL: AF106086; AAD47193.1; -
 DR EMBL: AB023045; BAB01716.1; -
 DR HSSP: P08659; 1LC1.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP-BINDING; 1.
 KM LIGase: Phenylpropanoid metabolism: Multigene family.
 FT CONFLICT 247 W -> L (IN REF. 2).
 FT CONFLICT 265 I -> V (IN REF. 2).
 SQ SEQUENCE 556 AA; 60929 MW; E3C2EAD677529EED CRC64;

Query Match 25.2%; Score 712; DB 1; Length 556;
 Best Local Similarity 34.8%; Pred. No. 3.4e-42;

Matches 178; Conservative 91; Mismatches 204; Indels 38; Gaps 8;

QY 46 NMLTGVDTYAEYLEKSCCLGEALKNYGLVVGRIALCSENCEFFPIVLAGLIGVGA 105

Db 57 NGPTGEVYTTADVAVTSRKLAAGLHNLGVKOHQNDVWMLLPNSPEVLLTEPLASFGAIRT 116
 QY 106 PTNEIYTLRELVLHSLGSKPTIVFSSKKGIDKVTYQK----TVT-----AIRTIVLDSK 157
 Db 117 SANDEFTPAEISKQAKASAKALVITQSRVYDKIKNLQNDGVLIYTTDSALPENCILRSE 176
 QY 158 VDRGYOSMDNFIKKNTPOGFKSGSKFYEVNKRQVALIMNSGSGTGLPKGVOLTHENI 217
 Db 177 LTQSEPRVDSIEKISP-----EDVALPFSSTGTGLPKGVMLTHKGL 220
 QY 218 VTRFSHARD-----PIYGNQVSPETALITVPPFHGGMFT-TLGYITGCFRIVMLTKPDE 272
 Db 221 VTSVAQVDGENNLFNR----DDVILCVMPHFHYALNSIMCSLRIGATLLIMPKFEI 277
 QY 273 ETEFLTLQDYKCSSVILVPTLFAILNRSELDKYDLSNVEIASGAPLKEIGEAVAR 332
 Db 278 TLLLEQIQCKKVAVAMVPPVILAIAKSPETEKEYDLSVBMVKSAGAPLGRLEDAISAK 337
 QY 333 FNLPGVROGUGLETTSALITTPEGDKP-----GASGVVPLFAKVIDLDTKKTLPGN 387
 Db 338 FPNAKLGQGYGMEAGPVLAMSLGFAKEFPYKSGAGCTVAVNNAEMKILDPDPTGSLPRN 397
 QY 388 RRGEEVCVKPMLMKGYVDNPEATREIIDEQMLHNGDIGYDEEKNHFTYDRLKSLITYK 447
 Db 398 KPEGLTIRNOQIMKGYLNDPLATASTIDKQMLHTGDVGFIDDDDELFTYDRLKELITYK 457
 QY 448 GIOVPPAELESVLLQHPNIFDAGVAGVDPPIAGELPGAVVYLKKGSKMTEKEMDYVASQ 507
 Db 458 GQVAPAELESLLIGHPEINDVAVAMKEDEAGEVPVAFVVASKSNISSEDEIKQFVSKQ 517
 QY 508 VSNAKRLRGVRFVDEVPKGLTGKIDGKAIR 538
 Db 518 VVFYKRI-NKVFETESIPKAPSGKILRKDLR 547

Search completed: November 29, 2001, 02:00:07
 Job time: 377 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2001, 02:00:07 ; Search time 25.81 Seconds

(without alignments)
778.471 Million cell updates/sec

Title: US-09-581-241-6

Perfect score: 2823
Sequence: 1 MENNENDENIYVGEPEPPYPI.....TGKIDGKAIKREILKKPKAKM 548

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SWISSPROT_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2814	99.7	548	1 LUCI_LUCCLA	Q01158 luciola lat
2	2684	95.1	548	1 LUCI_LUCCR	P13129 luciola cru
3	1945.5	68.9	550	1 LUCI_PPOPY	P08659 photinus py
4	1766.5	27.2	544	1 LUCI_PETCR	P14912 petroselinu
5	1758.5	26.9	544	1 LUCI_PETCR	P14913 petroselinu
6	1756.5	26.8	547	1 LUCI_TOBAC	O24145 nicotiana t
7	1746.5	26.4	542	1 LUCI_TOBAC	O24146 nicotiana t
8	1745	26.4	553	1 LUCI_VANPL	P31684 solanum tub
9	1732.5	25.9	545	1 LUCI_SOLTU	O24540 vanilla pla
10	1731	25.9	561	1 LUCI_ORYSA	O98777 arabidopsis
11	1730	25.9	563	1 LUCI_ORYSA	P17814 oryza sativ
12	1725.5	25.7	545	1 LUCI_SOLTU	P31685 solanum tub
13	1724	25.6	562	1 LUCI_ARYBN	P31687 glycine max
14	1718.5	25.5	561	1 LUCI_ARYBN	P31687 glycine max
15	1710	25.2	556	1 LUCI_ARYBN	O42524 arabidopsis
16	1691.5	24.5	569	1 LUCI_ORYSA	O98777 arabidopsis
17	1672	23.8	537	1 LUCI_PINTFA	P41636 pinus taeda
18	1587.5	20.8	293	1 LUCI_SOYBN	P31686 glycine max
19	1574.5	20.4	560	1 LUCI_BACSU	P94547 bacillus su
20	1568	20.1	562	1 LUCI_BACSU	P94547 bacillus su
21	1483	17.1	562	1 LUCI_HAEIN	P29212 escherichia
22	1465	16.5	543	1 LUCI_YEAST	P46510 haemophilus
23	1460	16.3	572	1 LUCI_BACSU	P31687 glycine max
24	1446	15.8	548	1 LUCI_BACSU	P31687 glycine max
25	1425.5	15.1	522	1 LUCI_ECOLI	P31687 glycine max
26	1370	13.1	698	1 LUCI_HUMAN	P31687 glycine max
27	1368	13.0	699	1 LUCI_MOUSE	P31687 glycine max
28	1361	12.8	699	1 LUCI_MOUSE	P31687 glycine max
29	1357	12.6	492	1 LUCI_MOUSE	P31687 glycine max
30	1353.5	12.5	699	1 LUCI_MOUSE	P31687 glycine max
31	1352	12.3	699	1 LUCI_MOUSE	P31687 glycine max
32	1348.5	12.3	699	1 LUCI_MOUSE	P31687 glycine max
33	1340	12.0	698	1 LUCI_MOUSE	P31687 glycine max

34	334.5	11.8	539	1 DHEB_BACSU	P40871 bacillus su
35	323	11.4	697	1 LUCI_HUMAN	O9ku0 homo sapien
36	323	11.4	711	1 LUCI_HUMAN	O60488 homo sapien
37	322.5	11.4	670	1 LUCI_RAT	O35547 rattus norv
38	321.5	11.4	6359	1 BACQ_BACLI	O68008 b bacitraci
39	319.5	11.3	694	1 LUCI_YEAST	P47912 saccharomyc
40	314	11.1	697	1 LUCI_YEAST	P31124 rattus norv
41	311.5	11.0	543	1 VIBE_VIBCH	O07899 vibrio chol
42	310.5	11.0	744	1 LUCI_YEAST	P39518 saccharomyc
43	307.5	10.9	744	1 LUCI_YEAST	P30624 saccharomyc
44	306.5	10.9	711	1 LUCI_YEAST	O9qu17 mus musculu
45	304	10.8	5255	1 BACA_BACLI	O68006 b bacitraci

ALIGNMENTS

RESULT	ID	Sequence	Score	Length	DB	Description
1	LUCI_LUCCLA	001158: luciola lat	2814	548	1	Q01158 luciola lat
2	LUCI_LUCCR	001159: luciola cru	2684	548	1	P13129 luciola cru
3	LUCI_PPOPY	008659: photinus py	1945.5	550	1	P08659 photinus py
4	LUCI_PETCR	014912: petroselinu	1766.5	544	1	P14912 petroselinu
5	LUCI_PETCR	014913: petroselinu	1758.5	544	1	P14913 petroselinu
6	LUCI_TOBAC	024145: nicotiana t	1756.5	547	1	O24145 nicotiana t
7	LUCI_TOBAC	024146: nicotiana t	1746.5	542	1	O24146 nicotiana t
8	LUCI_VANPL	024540: vanilla pla	1745	553	1	P31684 solanum tub
9	LUCI_SOLTU	098777: arabidopsis	1732.5	545	1	O98777 arabidopsis
10	LUCI_ORYSA	017814: oryza sativ	1731	561	1	P17814 oryza sativ
11	LUCI_SOLTU	031685: solanum tub	1730	563	1	P31685 solanum tub
12	LUCI_ARYBN	042524: arabidopsis	1725.5	545	1	P31687 glycine max
13	LUCI_ARYBN	042524: arabidopsis	1724	562	1	P31687 glycine max
14	LUCI_ORYSA	098777: arabidopsis	1718.5	561	1	O98777 arabidopsis
15	LUCI_ORYSA	098777: arabidopsis	1710	569	1	O98777 arabidopsis
16	LUCI_ORYSA	098777: arabidopsis	1691.5	569	1	O98777 arabidopsis
17	LUCI_PINTFA	041636: pinus taeda	1672	537	1	P41636 pinus taeda
18	LUCI_SOYBN	031686: glycine max	1587.5	293	1	P31686 glycine max
19	LUCI_BACSU	094547: bacillus su	1574.5	560	1	P94547 bacillus su
20	LUCI_BACSU	094547: bacillus su	1568	562	1	P94547 bacillus su
21	LUCI_HAEIN	029212: escherichia	1483	562	1	P29212 escherichia
22	LUCI_YEAST	046510: haemophilus	1465	543	1	P46510 haemophilus
23	LUCI_BACSU	031687: glycine max	1460	572	1	P31687 glycine max
24	LUCI_BACSU	031687: glycine max	1446	548	1	P31687 glycine max
25	LUCI_ECOLI	031687: glycine max	1425.5	522	1	P31687 glycine max
26	LUCI_HUMAN	031687: glycine max	1370	698	1	P31687 glycine max
27	LUCI_MOUSE	031687: glycine max	1368	699	1	P31687 glycine max
28	LUCI_MOUSE	031687: glycine max	1361	699	1	P31687 glycine max
29	LUCI_MOUSE	031687: glycine max	1357	492	1	P31687 glycine max
30	LUCI_MOUSE	031687: glycine max	1353.5	699	1	P31687 glycine max
31	LUCI_MOUSE	031687: glycine max	1352	699	1	P31687 glycine max
32	LUCI_MOUSE	031687: glycine max	1348.5	699	1	P31687 glycine max
33	LUCI_MOUSE	031687: glycine max	1340	698	1	P31687 glycine max

RESULT	ID	Sequence	Score	Length	DB	Description
1	LUCI_LUCCLA	001158: luciola lat	2814	548	1	Q01158 luciola lat
2	LUCI_LUCCR	001159: luciola cru	2684	548	1	P13129 luciola cru
3	LUCI_PPOPY	008659: photinus py	1945.5	550	1	P08659 photinus py
4	LUCI_PETCR	014912: petroselinu	1766.5	544	1	P14912 petroselinu
5	LUCI_PETCR	014913: petroselinu	1758.5	544	1	P14913 petroselinu
6	LUCI_TOBAC	024145: nicotiana t	1756.5	547	1	O24145 nicotiana t
7	LUCI_TOBAC	024146: nicotiana t	1746.5	542	1	O24146 nicotiana t
8	LUCI_VANPL	024540: vanilla pla	1745	553	1	P31684 solanum tub
9	LUCI_SOLTU	098777: arabidopsis	1732.5	545	1	O98777 arabidopsis
10	LUCI_ORYSA	017814: oryza sativ	1731	561	1	P17814 oryza sativ
11	LUCI_SOLTU	031685: solanum tub	1730	563	1	P31685 solanum tub
12	LUCI_ARYBN	042524: arabidopsis	1725.5	545	1	P31687 glycine max
13	LUCI_ARYBN	042524: arabidopsis	1724	562	1	P31687 glycine max
14	LUCI_ORYSA	098777: arabidopsis	1718.5	561	1	O98777 arabidopsis
15	LUCI_ORYSA	098777: arabidopsis	1710	569	1	O98777 arabidopsis
16	LUCI_ORYSA	098777: arabidopsis	1691.5	569	1	O98777 arabidopsis
17	LUCI_PINTFA	041636: pinus taeda	1672	537	1	P41636 pinus taeda
18	LUCI_SOYBN	031686: glycine max	1587.5	293	1	P31686 glycine max
19	LUCI_BACSU	094547: bacillus su	1574.5	560	1	P94547 bacillus su
20	LUCI_BACSU	094547: bacillus su	1568	562	1	P94547 bacillus su
21	LUCI_HAEIN	029212: escherichia	1483	562	1	P29212 escherichia
22	LUCI_YEAST	046510: haemophilus	1465	543	1	P46510 haemophilus
23	LUCI_BACSU	031687: glycine max	1460	572	1	P31687 glycine max
24	LUCI_BACSU	031687: glycine max	1446	548	1	P31687 glycine max
25	LUCI_ECOLI	031687: glycine max	1425.5	522	1	P31687 glycine max
26	LUCI_HUMAN	031687: glycine max	1370	698	1	P31687 glycine max
27	LUCI_MOUSE	031687: glycine max	1368	699	1	P31687 glycine max
28	LUCI_MOUSE	031687: glycine max	1361	699	1	P31687 glycine max
29	LUCI_MOUSE	031687: glycine max	1357	492	1	P31687 glycine max
30	LUCI_MOUSE	031687: glycine max	1353.5	699	1	P31687 glycine max
31	LUCI_MOUSE	031687: glycine max	1352	699	1	P31687 glycine max
32	LUCI_MOUSE	031687: glycine max	1348.5	699	1	P31687 glycine max
33	LUCI_MOUSE	031687: glycine max	1340	698	1	P31687 glycine max

RESULT	ID	Sequence	Score	Length	DB	Description
1	LUCI_LUCCLA	001158: luciola lat	2814	548	1	Q01158 luciola lat
2	LUCI_LUCCR	001159: luciola cru	2684	548	1	P13129 luciola cru
3	LUCI_PPOPY	008659: photinus py	1945.5	550	1	P08659 photinus py
4	LUCI_PETCR	014912: petroselinu	1766.5	544	1	P14912 petroselinu
5	LUCI_PETCR	014913: petroselinu	1758.5	544	1	P14913 petroselinu
6	LUCI_TOBAC	024145: nicotiana t	1756.5	547	1	O24145 nicotiana t
7	LUCI_TOBAC	024146: nicotiana t	1746.5	542	1	O24146 nicotiana t
8	LUCI_VANPL	024540: vanilla pla	1745	553	1	P31684 solanum tub
9	LUCI_SOLTU	098777: arabidopsis	1732.5	545	1	O98777 arabidopsis
10	LUCI_ORYSA	017814: oryza sativ	1731	561	1	P17814 oryza sativ
11	LUCI_SOLTU	031685: solanum tub	1730	563	1	P31685 solanum tub
12	LUCI_ARYBN	042524: arabidopsis	1725.5	545	1	P31687 glycine max
13	LUCI_ARYBN	042524: arabidopsis	1724	562	1	P31687 glycine max
14	LUCI_ORYSA	098777: arabidopsis	1718.5	561	1	O98777 arabidopsis
15	LUCI_ORYSA	098777: arabidopsis	1710	569	1	O98777 arabidopsis
16	LUCI_ORYSA	098777: arabidopsis	1691.5	569	1	O98777 arabidopsis
17	LUCI_PINTFA	041636: pinus taeda	1672	537	1	P41636 pinus taeda
18	LUCI_SOYBN	031686: glycine max	1587.5	293	1	P31686 glycine max
19	LUCI_BACSU	094547: bacillus su	1574.5	560	1	P94547 bacillus su
20	LUCI_BACSU	094547: bacillus su	1568	562	1	P94547 bacillus su
21	LUCI_HAEIN	029212: escherichia	1483	562	1	P29212 escherichia
22	LUCI_YEAST	046510: haemophilus	1465	543	1	P46510 haemophilus
23	LUCI_BACSU	031687: glycine max	1460	572	1	P31687 glycine max
24	LUCI_BACSU	031687: glycine max	1446	548	1	P31687 glycine max
25	LUCI_ECOLI	031687: glycine max	1425.5	522	1	P31687 glycine max
26	LUCI_HUMAN	031687: glycine max	1370	698	1	P31687 glycine max
27	LUCI_MOUSE	031687: glycine max	1368	699	1	P31687 glycine max
28	LUCI_MOUSE	031687: glycine max	1361	699	1	P31687 glycine max
29	LUCI_MOUSE	031687: glycine max	1357	492	1	P31687 glycine max
30	LUCI_MOUSE	031687: glycine max	1353.5	699	1	P31687 glycine max
31	LUCI_MOUSE	031687: glycine max	1352	699	1	P31687 glycine max
32	LUCI_MOUSE	031687: glycine max	1348.5	699	1	P31687 glycine max
33	LUCI_MOUSE	031687: glycine max	1340	698	1	P31687 glycine max

RESULT	1				
LUCI_LUCCLA					
ID	LUCI_LUCCLA	STANDARD;	PRT:	548 AA.	
AC	001158;				
DT	01-APR-1993 (rel. 25, Created)				
DT	01-APR-1993 (rel. 25, Last sequence update)				
DT	20-AUG-2001 (rel. 40, Last annotation update)				
DE	LUCIFERIN 4-MONOOXYGENASE (EC 1.13.12.7) (LUCIFERASE).				
OS	Luciola lateralis (firefly).				
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
OC	Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;				
OC	Eletheroformia; Cantharoidea; Lampyridae; Luciola.				
OX	NCBI_TaxID=7052;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92305054; PubMed=1610896;				
RA	Tatsumi H., Kajiyama N., Nakano E.;				
RT	"Molecular cloning and expression in Escherichia coli of a cDNA clone				
RT	encoding luciferase of a firefly, Luciola lateralis.";				
RL	Biochim. Biophys. Acta 1131:161-165(1992).				
CC	-1- FUNCTION: PRODUCES GREEN LIGHT.				
CC	-1- CATALYTIC ACTIVITY: LUCIFERIN + O(2) + ATP = OXIDIZED LUCIFERIN +				
CC	CO(2) + H(2)O + AMP + PYROPHOSPHATE + LIGHT.				
CC	-1- COFACTOR: REQUIRES MAGNESIUM.				
CC	-1- SUBCELLULAR LOCATION: PEROXISOME.				
CC	-1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME				
CC	FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on ways				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
CC	EMBL; X66919; CAA47358.1; -.				
DR	PIR: S23437; S23437.				
DR	HSSP: P08659; ILCI.				
DR	InterPro: IPR000873; AMP-bind.				
DR	Pfam: PF00501; AMP-binding; 1.				
DR	PROSITE: PS00342; MICROBODIES_CTER; 1.				
DR	PROSITE: PS00455; AMP-BINDING; 1.				
KW	Oxidoreductase; Monooxygenase; Photoprotein; Luminescence; Magnesium;				
KW	Peroxisome.				
FT	SITE	546	548	MICROBODY TARGETING SIGNAL (POTENTIAL).	
SO	SEQUENCE	548 AA;	60125 MW;	ACG2FP3320BBDDA46 CRC64;	
	Query Match	99.7%;	Score 2814;	DB 1;	Length 548;
	Best Local Similarity	99.6%;	Pred. No. 2,86-188;		
	Matches 546;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;
07	1	MENNENDENIVGPEPPYPIEESGAGOLRKMDYATLGAIAATNALTVGDTVAEYLE	60		
DB	1	MENNENDENIVGPEPPYPIEESGAGOLRKMDYATLGAIAATNALTVGDTVAEYLE	60		

```

QY 61 KSCCIGALKNGVLYVDGRIALCSENCEEFFIPVLAGEIFGVGVAFTNFIYTLRELHVSU 120
DB 61 KSCCIGALKNGVLYVDGRIALCSENCEEFFIPVLAGEIFGVGVAFTNFIYTLRELHVSU 120
QY 121 GISKPTIYSSSKGKGLDKVTVVOKTATYATITVIILDSKVDYRGYOSMDNFIKNTPOGFGK 180
DB 121 GISKPTIYSSSKGKGLDKVTVVOKTATYATITVIILDSKVDYRGYOSMDNFIKNTPOGFGK 180
QY 181 SSFKTVEVNRKQVALIMNSSGSTGLPKGVOLTHENIYVRFSHARDPIYGNVSPGTAL 240
DB 181 SSFKTVEVNRKQVALIMNSSGSTGLPKGVOLTHENIYVRFSHARDPIYGNVSPGTAL 240
QY 241 TVPPEHGFEMFTTGLYLCGFRIVMLTKFDETEFLKTLQDYKCSSVILVPLFALINRS 300
DB 241 TVPPEHGFEMFTTGLYLCGFRIVMLTKFDETEFLKTLQDYKCSSVILVPLFALINRS 300
QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAARFNLPGVROGGLTETTSALITTEPGDCK 360
DB 301 ELLDKYDLSNLVEIASGAPLSKEIGEAARFNLPGVROGGLTETTSALITTEPGDCK 360
QY 361 PGASGVVPLFKAKYIDIDTKKTLGPNRRGEVCVGMIMKGYVNPPEATREIIDEGBUL 420
DB 361 PGASGVVPLFKAKYIDIDTKKTLGPNRRGEVCVGMIMKGYVNPPEATREIIDEGBUL 420
QY 421 HTGDIGYDEEKNHFTVDRLSLTKYKGQVPPAELESVLLQHPNIPDAGVAGVDPPIAG 480
DB 421 HTGDIGYDEEKNHFTVDRLSLTKYKGQVPPAELESVLLQHPNIPDAGVAGVDPPIAG 480
QY 481 ELPGAVVVLKKGKSMTEKEVMDYVASOVSNARKLRGVRFVDEVRKGLTKIDGKAIREI 540
DB 481 ELPGAVVVLKKGKSMTEKEVMDYVASOVSNARKLRGVRFVDEVRKGLTKIDGKAIREI 540
QY 541 LKKPVAKM 548
DB 541 LKKPVAKM 548

RESULT 2
LUC1_LUCCR STANDARD; PRT: 548 AA.
ID LUC1_LUCCR STANDARD; PRT: 548 AA.
AC P13129;
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE LUCIFERIN 4-MONOOXYGENASE (EC 1.13.12.7) (LUCIFERASE).
OS Luciola cruciata (Japanese firefly) (Genji firefly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Elateriformia; Cantharoidae; Lampyridae; Luciola.
OX NCBI_TaxID=7051;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=69326143; PubMed=2473944;
RA Masuda T., Tatsuji H., Nakano E.;
RT "Cloning and sequence analysis of cDNA for luciferase of a Japanese
RL firefly, Luciola cruciata."
CC -1- FUNCTION: PRODUCES GREEN LIGHT WITH A WAVELENGTH OF 544 NM.
CC -1- CATALYTIC ACTIVITY: LUCIFERIN + O(2) + ATP = OXIDIZED LUCIFERIN +
CC CO(2) + H(2)O + AMP + PYROPHOSPHATE + LIGHT.
CC -1- COFACTOR: REQUIRES MAGNESIUM.
CC -1- SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
CC EMBL: M26194; AAA29135.1;
CC PIR: J50181; J50181.
CC HSP: P08659; 11C1.
CC InterPro: IPR000873; AMP-bind.
CC Pfam: PF00501; AMP-binding; 1.
CC ProSite: PS00342; MICROBODIES_CTER; 1.
CC ProSite: PS00455; AMP BINDING; 1.
CC Oxidoreductase; Monooxygenase; Photoprotein; Luminescence; Magnesium;
CC Pteroisome.
CC SITE 546 548 MICROBODY TARGETING SIGNAL (POTENTIAL).
CC FT SEQUENCE 548 AA: 60017 MW: 2052D6189E79109F CRC64.
CC SQ

Query Match          95.1%; Score 2684; DB 1: Length 548;
Best Local Similarity 93.4%; Pred. No. 3e-179;
Matches 512; Conservative 26; Mismatches 10; Indels 0; Gaps 0;

QY 1 MENMENDENIYVGPPEFYDIEGSAQOLRKYMRYAKIGAIATNALGVDTYAYEYLE 60
DB 1 MENMENDENIYVGPPEFYDIEGSAQOLRKYMRYAKIGAIATNALGVDTYAYEYLE 60
QY 61 KSCCIGALKNGVLYVDGRIALCSENCEEFFIPVLAGEIFGVGVAFTNFIYTLRELHVSU 120
DB 61 KSCCIGALKNGVLYVDGRIALCSENCEEFFIPVLAGEIFGVGVAFTNFIYTLRELHVSU 120
QY 121 GISKPTIYSSSKGKGLDKVTVVOKTATYATITVIILDSKVDYRGYOSMDNFIKNTPOGFGK 180
DB 121 GISKPTIYSSSKGKGLDKVTVVOKTATYATITVIILDSKVDYRGYOSMDNFIKNTPOGFGK 180
QY 181 SSFKTVEVNRKQVALIMNSSGSTGLPKGVOLTHENIYVRFSHARDPIYGNVSPGTAL 240
DB 181 SSFKTVEVNRKQVALIMNSSGSTGLPKGVOLTHENIYVRFSHARDPIYGNVSPGTAL 240
QY 241 TVPPEHGFEMFTTGLYLCGFRIVMLTKFDETEFLKTLQDYKCSSVILVPLFALINRS 300
DB 241 TVPPEHGFEMFTTGLYLCGFRIVMLTKFDETEFLKTLQDYKCSSVILVPLFALINRS 300
QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAARFNLPGVROGGLTETTSALITTEPGDCK 360
DB 301 ELLDKYDLSNLVEIASGAPLSKEIGEAARFNLPGVROGGLTETTSALITTEPGDCK 360
QY 361 PGASGVVPLFKAKYIDIDTKKTLGPNRRGEVCVGMIMKGYVNPPEATREIIDEGBUL 420
DB 361 PGASGVVPLFKAKYIDIDTKKTLGPNRRGEVCVGMIMKGYVNPPEATREIIDEGBUL 420
QY 421 HTGDIGYDEEKNHFTVDRLSLTKYKGQVPPAELESVLLQHPNIPDAGVAGVDPPIAG 480
DB 421 HTGDIGYDEEKNHFTVDRLSLTKYKGQVPPAELESVLLQHPNIPDAGVAGVDPPIAG 480
QY 481 ELPGAVVVLKKGKSMTEKEVMDYVASOVSNARKLRGVRFVDEVRKGLTKIDGKAIREI 540
DB 481 ELPGAVVVLKKGKSMTEKEVMDYVASOVSNARKLRGVRFVDEVRKGLTKIDGKAIREI 540
QY 541 LKKPVAKM 548
DB 541 LKKPVAKM 548

RESULT 3
LUC1_PROPY STANDARD; PRT: 550 AA.
ID LUC1_PROPY STANDARD; PRT: 550 AA.
AC P08659;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE LUCIFERIN 4-MONOOXYGENASE (EC 1.13.12.7) (LUCIFERASE).
OS Photinus pyralis (North American firefly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Elateriformia; Cantharoidae; Lampyridae; Photinus.
OX NCBI_TaxID=7054;
RN [1]

```

RA SEQUENCE FROM N.A. PubMed=3821727.
 RX MEDLINE=87144243; PubMed=3821727.
 RA de Wet J.R., Wood K.V., Deluca M., Hellinski D.R., Subramani S.;
 RT "Firefly Luciferase gene: structure and expression in mammalian
 cells.";
 RT Mol. Cell. Biol. 7:725-737(1987).
 RL [2]
 RN SUBCELLULAR LOCATION.
 RP MEDLINE=87204117; PubMed=3554235;
 RX Kellier G.-A., Gould S., de Luca M., Subramani S.;
 RA "Firefly luciferase is targeted to peroxisomes in mammalian cells.";
 RT Proc. Natl. Acad. Sci. U.S.A. 84:3264-3268(1987).
 RL [3]
 RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RP MEDLINE=96398615; PubMed=8805533;
 RX Conti E., Franks N.P., Brick P.;
 RA "Crystal structure of firefly luciferase throws light on a
 RT superfamily of adenylate-forming enzymes.";
 RL Structure 4:287-298(1996).
 RN [4]
 RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RP MEDLINE=99007339; PubMed=9788915;
 RX Franks N.P., Jenkins A., Conti E., Lieb W.R., Brick P.;
 RA "Structural basis for the inhibition of firefly luciferase by a
 RT general anesthetic.";
 RL Biophys. J. 75:2205-2211(1998).
 CC -1- FUNCTION: PRODUCES GREEN LIGHT WITH A WAVELENGTH OF 562 NM.
 CC -1- CATALYTIC ACTIVITY: LUCIFERIN + O(2) + ATP = OXIDIZED LUCIFERIN +
 CC CO(2) + H(2O) + AMP + PYROPHOSPHATE + LIGHT.
 CC -1- COFACTOR: REQUIRES MAGNESIUM.
 CC -1- SUBCELLULAR LOCATION: PEROXISOML.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M15077; AAA29795.1; -;
 CC EMBL: X84848; CAAS9283.1; -;
 CC EMBL: U03687; AAA03561.1; -;
 CC EMBL: U09934; AAB64396.1; -;
 CC EMBL: U89935; AAB64399.1; -;
 CC PIR: A26772; A26772.
 CC PDB: 1LCI; 26-MAR-97.
 CC PDB: 1BA3; 11-NOV-98.
 CC InterPro: IPR000873; AMP-bind.
 CC Pfam: PF00501; AMP-binding; 1.
 CC PROSITE: PS00342; MICROBODIES_CTER; 1.
 CC PROSITE: PS00455; AMP_BINDING; 1.
 CC Oxidoreductase; Monooxygenase; Photoprotein; Luminescence; Magnesium;
 KW Peroxisome; 3D-structure.
 KW SITE 548 550 MICROBODY TARGETING SIGNAL.
 SQ SEQUENCE 550 AA; 60745 MW; E380FC9E956ACD6 CRC64;

Db	121	SQPTVFEVSKKGLQKILINVOKKLPITOKIILIMOSKTYQGFQSMYTFVTVSHLPPEHND	180
QY	183	FKTEVNRKEQEOVALIMNMSGSTGLPKGVOLTENENIYTRFSHARDPIYGNVSPGTALITV	242
Db	181	FVPSFPRDCKITALIMNMSGSTGLPKGVALPHRIACRFSHARPIFGNOIIPDFTALSV	240
QY	243	VPEHHGEMPTTLCYLTGCFRIYMLTFPDEETFLKTIQDYKCSSVILVPLFALINSEL	302
Db	241	VPFHNGEMPTTLCYLTGCFRVYVLMYFEEELPLRSIDQYKISALLVPLFSFFAKSTL	300
QY	303	LDKVDLSNLVLEIASGGAPLSKEIEEAAARRNLPBGVQNGGLTETTSALIIITPBGDKPG	362
Db	301	IDKVDLSNLHLEIASGGAPLSKEVEEAAAKRPHLPGIKQGGLETETTSALIIITPBGDKPG	360
QY	363	ASGVVPLFLFKKVIDLDTKKTIGNRGGEVCKPMLKMGVNDPEATREITDEEGLHT	422
Db	361	AVGVVLPFFFEKAVDDLTGKTLGVNQGELCYVRKPMISGTVNNPEATNALIDKDGHLHS	420
QY	423	GDIGYDEEKHEFFIVDRLSLIDIKKGYQVPAELDESVLQHPNIFDAGVADPPIAGEL	482
Db	421	GDIAVWDEDEHFFIVDRILKSILTKKGYQVAPAELESTLLQHPNIFDAGVAGLPDDOAGEL	480
QY	483	PGAVVVLKKKGSMTKEVMDIVASQVSNAKRLRGVRPVDEVPKGLTGKIDGAKIREILK	542
Db	481	PAVVVLEHNGKTKTEKEITVDVAVASQVTTAKKLRGCVVFEVDEVPKGLTGKIDAKIREILI	540
QY	543	K 543	
Db	541	K 541	

[illegible]

```

CC RESULT 4
CC ACCL_PENCOR
CC ID ACCL_PENCOR STANDARD: PRT: 544 AA.
CC AC P14912.
CC DT 01-APR-1990 (Rel. 14, Created)
CC DT 01-APR-1990 (Rel. 14, Last sequence update)
CC DT 20-AUG-2001 (Rel. 40, Last annotation update)
CC DE 4-COMMARATE--COA LIGASE 1 (EC 6.2.1.12) (ACCL 1) (4-COMMAROYL-COA
CC PHENYLPIRUVATE 1).
CC CN ACCL OR ACCL-1.
CC OS Petroselinum crispum (Parsley) (Petroselinum hortense).
CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
CC OC Asteridae; euasterids II; Apiales; Apiaceae; Petroselinum.
CC OX NCBI_TaxID=4043;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=89005119; PubMed=3169018;
CC RA Lozoya E., Hoffmann H., Douglas C., Schulz W., Scheel D.,
CC RA Hahlbrock K.;
CC RT "Primary structures and catalytic properties of isoenzymes encoded by
CC RT the two 4-commarate:COA ligase genes in parsley.";
CC RL Eur. J. Biochem. 176:661-667(1988).
CC RN [2]
CC RP SEQUENCE OF 1-8 FROM N.A.
CC RA Douglas C., Hoffmann H., Schulz W., Hahlbrock K.;
CC RT "Structure and elicitor or U.V.-light-stimulated expression of two
CC RT 4-commarate:COA ligase genes in parsley.";
CC RL EMBE J. 6:1189-1195(1987).
CC CC -1- CATALYTIC ACTIVITY: ATP + 4-COMMARATE + COA = AMP + DIPHOSPHATE +
CC CC 4-COMMAROYL-COA.
CC CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROANOID
CC CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
CC CC -1- INDUCTION: TRANSCRIPTIONALLY ACTIVATED BY BOTH FUNGAL ELICTOR AND
CC CC ULTRAVIOLET IRRADIATION.
CC CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC CC FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation at
CC CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
CC or send an email to license@isb-sdb.ch).

DR EMBL, X13324; CAA31696.1; -
DR EMBL, X05350; CAA28959.1; -
DR PIR, S01667; S01667.
DR HSSP, P08659; 11C1.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE, PS00453; AMP_BINDING; 1.
DR Ligase, Phenylpropanoid metabolism; Multigene family.
KW SEQUENCE 544 AA; 59825 MW; 22BBAD78F255D0C8 CRC64;

Query Match 27.2%; Score 766.5; DB 1; Length 544;
Best Local Similarity 36.0%; Pred. No. 6.3e-46;
Matches 186; Conservative 89; Mismatches 189; Indels 53; Gaps 10;

QY 46 NALTGVYVYAEYLEKSCCLGALKNYGLVNDGRIALCSENEEFPFIVLAGLFGVGA 105
DB 47 NCATGETFYVSVVLLSKRVASGLKLGIOGDTIMLLPNSEYFFALGASYRGAIST 106
QY 106 PTNEYTYTRELVLHSLGSKPTVFSKKGDLRVTVQKTVAIKTVIL--DSKVDYRG 162
DB 107 MANPFTSAEVYIKOLKASQAKLITQACVYDKV---KYAAEKNIQIICIDDAPODCH 162
QY 163 YOSMNFIKKNTPOGFGKSSFTVEYNREKQVALIMNSSGSGTGLPKGVOLTHENIVTR-- 220
DB 163 FSKLMEADSEMP-----VVINSDDVAL--PYSSGTGLPKGVMLTHKGLVTSVA 212
QY 221 -----FSHARDPIYGNQVSPGTALIVVPHHGGMFTTGLYLCGRF---IYM 266
DB 213 QQVGDNDPNLYMHSD-----VMICLPLFH---IYSLANVLCCLGRLAGVTILI 258
QY 267 LTRPEETFLKTLQDYKCSSVILVPTLFAILNRSELLDYDLSNVEIASGAPLSKEIG 326
DB 259 MOKFIVLPTELILQKXYKVIIGFVPPYVIALAKSPVVDYDSSRYTVWSGAPLKELE 318
QY 327 EAVARFNLPGVRGQYGLTETTSATII-----TPGDDKPGASGVVPLFKAKVIDLDTK 381
DB 319 DAVRAKFPNAKIGQGYGMEAGPVTLAMCLAFAPKEPEIKSGAGTVVRAEMKIVDPETN 378
QY 382 KTLGNRRGVEVCKPMLKGYVNDPEATRELIIDEGWLHTDIDGYDEKHEFFIVDRK 441
DB 379 ASLPRNKGELICIRGDIQIMKGYLNDPESTRITIDEGWMLHTDIDGIDDELFTVDRK 438
QY 442 SLIRKGYVPAELESVYLQHPNIFDAGVAGVPPPIAGELPGAVVYLKKGSMTEKEVM 501
DB 439 ELIKKKGQVAAELALLHTPTISDAVVMIDEKAGEVAVFVVRINGFTTTEEEK 498
QY 502 DYVASOVNAKRLRGVRVDEVPKGLTKIDGKAIR 538
DB 499 QPVSKQVVFYKRI-FRVFFVDAIPKSPSGKILRKDR 534

RESULT 5
AC12_PETCR STANDARD: PRT: 544 AA.
ID AC12_PETCR
AC P14913;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE 4-COUMARATE--COA LIGASE 1 (EC 6.2.1.12) (4CL 1) (4-COUMAROYL-COA
DE SYNTHASE 1).
GN 4CL2 OR 4CL2.
OS Petroselinum crispum (Parsley) (Petroselinum hortense).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Apiales; Apiaceae; Petroselinum.
OC NCBI_TaxId=4043;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE-89005119; PubMed-3169018;
RA Lozoya E., Hoffmann H., Douglas C., Schulz W., Scheel D.,
RA Hahlbrock K.,
RT "Primary structures and catalytic properties of isoenzymes encoded by
RT the two 4-coumarate:COA ligase genes in parsley."
RL Eur. J. Biochem. 176:661-667(1988).
RN [2]
RP SEQUENCE OF 1-8 FROM N.A.
RA Douglas C., Hoffmann H., Schulz W., Hahlbrock K.,
RT "Structure and elicitor or U.V.-light-stimulated expression of two
RT 4-coumarate:COA ligase genes in parsley."
RL EMBO J. 6:1189-1195(1987).
CC -1- CATALYTIC ACTIVITY: ATP + 4-COUMARATE + COA = AMP + DIPHOSPHATE +
CC 4-COUMAROYL-COA.
CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
CC -1- INDUCTION: TRANSCRIPTIONALLY ACTIVATED BY BOTH FUNGAL ELICITOR AND
CC ULTRAVIOLET IRRADIATION.
CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
CC or send an email to license@isb-sdb.ch).

DR EMBL, X13325; CAA31697.1; -
DR EMBL, X05351; CAA28960.1; -
DR PIR, S01667; S01667.
DR PIR, S15695; S15695.
DR HSSP, P08659; 11C1.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE, PS00453; AMP_BINDING; 1.
KW Ligase, Phenylpropanoid metabolism; Multigene family.
KW SEQUENCE 544 AA; 59783 MW; B477965CC6F8C534 CRC64;

Query Match 26.9%; Score 758.5; DB 1; Length 544;
Best Local Similarity 35.8%; Pred. No. 2.3e-45;
Matches 185; Conservative 89; Mismatches 190; Indels 53; Gaps 10;

QY 46 NALTGVYVYAEYLEKSCCLGALKNYGLVNDGRIALCSENEEFPFIVLAGLFGVGA 105
DB 47 NCATGETFYVSVVLLSKRVASGLKLGIOGDTIMLLPNSEYFFALGASYRGAIST 106
QY 106 PTNEYTYTRELVLHSLGSKPTVFSKKGDLRVTVQKTVAIKTVIL--DSKVDYRG 162
DB 107 MANPFTSAEVYIKOLKASLAKLITQACVYDKV---KYAAEKNIQIICIDDAPODCH 162
QY 163 YOSMNFIKKNTPOGFGKSSFTVEYNREKQVALIMNSSGSGTGLPKGVOLTHENIVTR-- 220
DB 163 FSKLMEADSEMP-----VVINSDDVAL--PYSSGTGLPKGVMLTHKGLVTSVA 212
QY 221 -----FSHARDPIYGNQVSPGTALIVVPHHGGMFTTGLYLCGRF---IYM 266
DB 213 QQVGDNDPNLYMHSD-----VMICLPLFH---IYSLANVLCCLGRLAGVTILI 258
QY 267 LTRPEETFLKTLQDYKCSSVILVPTLFAILNRSELLDYDLSNVEIASGAPLSKEIG 326
DB 259 MOKFIVLPTELILQKXYKVIIGFVPPYVIALAKSPVVDYDSSRYTVWSGAPLKELE 318
QY 327 EAVARFNLPGVRGQYGLTETTSATII-----TPGDDKPGASGVVPLFKAKVIDLDTK 381
DB 319 DAVRAKFPNAKIGQGYGMEAGPVTLAMCLAFAPKEPEIKSGAGTVVRAEMKIVDPETN 378
QY 382 KTLGNRRGVEVCKPMLKGYVNDPEATRELIIDEGWLHTDIDGYDEKHEFFIVDRK 441
DB 379 ASLPRNKGELICIRGDIQIMKGYLNDPESTRITIDEGWMLHTDIDGIDDELFTVDRK 438

QY 442 SLIKKGYPPAPLESVILQHPNIPDAGVAGVDPPIAGLPGAVVYLKKGSMTEKEYM 501
 DB 439 ELIKKGYPPAPLEALLHTHTPISDAVAVPIDERAGVPAFVAVRTNGFTTEELK 498
 QY 502 DYVASQVSNARKRLRGVFEVDEPKGLTGKIDGKAIR 538
 DB 499 QFVSKQVVFYKRI-FRFFVFDALPKSPSKILKKDLR 534
 RESULT 6
 4CL2_TOBAC STANDARD; PRT; 547 AA.
 AC 024145;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE 4-COUMARATE--COA LIGASE 1 (EC 6.2.1.12) (4CL 1) (4-COUMAROYL-COA
 SYNTHASE 1).
 GN 4CL1.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OC NCBI_TaxId=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96416441; PubMed=8819324;
 RA Lee D., Douglas C.J.;
 RT "Two divergent members of a tobacco 4-coumarate:coenzyme A ligase
 (4CL) gene family. cDNA structure, gene inheritance and expression,
 RT and properties of recombinant proteins."
 RL Plant Physiol. 112:193-205(1996).
 CC -1- CATALYTIC ACTIVITY: ATP + 4-COUMARATE + COA = AMP + DIPHOSPHATE +
 CC 4-COUMAROYL-COA.
 CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
 CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U50845; AAB18637.1; -
 DR HSSP: P08659; IBA3.
 DR Mendel: 24477; Nicita:1179;24478.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP-BINDING; 1.
 KW Ligase; Phenylpropanoid metabolism; Multi-gene family.
 SQ SEQUENCE 547 AA; 59842 MW; 154DF6D684E3F51B CRC64;
 Query Match 26.8%; Score 756.5; DB 1; Length 547;
 Best Local Similarity 34.9%; Pred. No. 3.2e-45;
 Matches 188; Conservative 90; Mismatches 205; Indels 55; Gaps 12;
 QY 32 YMDRAKGAIAFNTN-----ALNGVD--YTVAEYLEKSC-CGEGALKNYGLVVDG 78
 DB 25 YIPKHLPHISCFENISEFSSRPLINGANDQITVAF-VELTRKRAVGLINKIGIOKD 83
 QY 79 RIALSCENCEEFPIVLAGLIGVAVPTEITVRELVHSLGSKPTIVFSSKKGIDKY 138
 DB 84 TIMILPSPPEFAFMGASVLAISTMANPLFPAEVVQAKASAKIITQSCFVGKV 143
 QY 139 ITVQKTVAIKTIYILDSKVYRGYSMDNFIKNTPOG-----FKGSFKEVEVN-R 190
 DB 144 -----KDYASENDVKYICIDSAPEGCIHFSELTQSDHEIPEVKIQ 184

QY 191 KEOVALINMSGSTGLPKGYOLTHENIVTRSHARDPIYGN-QVSPGAILTVVPEHHGF 249
 DB 185 PDDVALPYSSTGTGLPKVMLTHKGLVTSAQOYDGENANILYMSEVYLKCVLPFH-- 242
 QY 250 GMEFTLGLYTCGFR-----IVMLTKFDEETFLKTLQDYKCSSVILPVLFAILNRSELDK 305
 DB 243 -IYLSNLSILGLRGAAILIMQKFDIAFPELLIGKYVSGISGFEVPIVLAIAKSPYDS 301
 QY 306 YDLSNLVEIASGAPLSKEIGAVARRENPGVROGYGLTFTTSAIIT-----IPSGDK 360
 DB 302 YDLSNVRTWASGAPLGLKLEDAVTRKFPNALKGGYGTEGAPVLAACLAFAKKEPFIK 361
 QY 361 PGASGVVPLFAKAVYIDLTKKTLGNRRGECVAGPMIMKGYVNPATREIIDEGML 420
 DB 362 SGACGVVNAEMKIVDPPTGCSLPRNGEELICINGDQIMKCYLNDPATRTIDKEGML 421
 QY 421 HNGDIGYDEERHFFIVDLRLKSLIRKGYVPPAPLESVILQHPNIPDAGVAGVDPPIAG 480
 DB 422 HNGDIGFIDEDELFIYDLRLKELIKKGYVAPAEIEALLNHPNISDAVAVPMKDEQAG 481
 QY 481 ELPGAVVYLKKGKSMTEKEYMDYVASQVSNARKRLRGVFEVDEPKGLTGKIDGKAIR 538
 DB 482 EVPAFVAVRSNGSAITDEVKDFISKQVIFYKRVK-RVFEVTVPKSPSKILRRDLR 538
 RESULT 7
 4CL2_TOBAC STANDARD; PRT; 542 AA.
 AC 024146;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE 4-COUMARATE--COA LIGASE 2 (EC 6.2.1.12) (4CL 2) (4-COUMAROYL-COA
 SYNTHASE 2).
 GN 4CL2.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OC NCBI_TaxId=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96416441; PubMed=8819324;
 RA Lee D., Douglas C.J.;
 RT "Two divergent members of a tobacco 4-coumarate:coenzyme A ligase
 (4CL) gene family. cDNA structure, gene inheritance and expression,
 RT and properties of recombinant proteins."
 RL Plant Physiol. 112:193-205(1996).
 CC -1- CATALYTIC ACTIVITY: ATP + 4-COUMARATE + COA = AMP + DIPHOSPHATE +
 CC 4-COUMAROYL-COA.
 CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
 CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U50846; AAB18638.1; -
 DR HSSP: P08659; IBA3.
 DR Mendel: 24478; Nicita:1179;24478.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP-BINDING; 1.
 KW Ligase; Phenylpropanoid metabolism; Multi-gene family.
 SQ SEQUENCE 542 AA; 59479 MW; CB5579AEDFCF003 CRC64;


```

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91217100; PubMed=2022667;
RA Becker-Andre M., Schulze-Isferl P., Hahlbrock K.;
RT "Structural comparison, modes of expression, and putative cis-acting
RT elements of the two 4-coumarate: CoA ligase genes in potato.";
RL J. Biol. Chem. 266:8551-8559(1991)
CC -1- CATALYTIC ACTIVITY: ATP + 4-COUMARATE + COA = AMP + DIPHOSPHATE +
CC 4-COUMAROYL-COA.
CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M62755; AAA3842.1; -
CC PIR; A39827; A39827.
CC HSSP; P08659; ILCT.
CC InterPro: IPR000873; AMP-bind.
CC Pfam: PF00501; AMP-binding; 1.
CC PROSITE; PS00455; AMP-BINDING; 1.
CC Ligase; Phenylpropanoid metabolism; Multigene family.
KW SEQUENCE 545 AA; 59619 MW; DE183683B74BA7A CRL64;
SQ

Query Match 25.9%; Score 732.5; DB 1; Length 545;
Best Local Similarity 34.0%; Pred. No. 1.5e-43;
Matches 184; Conservative 94; Mismatches 202; Indels 61; Gaps 12;

OY 32 YMDRYAKLGAIFFTN-----ALTGVD--VTYAEYLEKSCCLGALKNGYLVNQR 79
DB 23 YIPKHLPLHSYCFENSEFNSRPLCDGANDITTYAEELTSKRVANGLKLTGQOKDT 82
OY 80 IALSCNCEEFITPVLAGLFIVGVAAPNEITYLRELVHSLGDISKPTIVFSSKGLDVT 139
DB 83 IMLLEDCPEFEVFAFGASYLGAISTMANPLFTPAEVVQAKASAKIYITQACPAKV- 141
OY 140 TVQKVTATKTIYIIDSQVDYRGYSKDNFK-----KNTPOG-----FKSSKTYEV 188
DB 142 -----KDY-----AIENDLKVICVDSPVPEGCVHFSELIQSDHEHPIDV 179
OY 189 N-RKEQVALINMSGSTGLPKGVOLTHTENIVRFSHARPIYGN-QVSPGTAILLVVPFH 246
DB 180 KIQPDVVALPSSSGTGLPKVMTLHGGLVTSVAQVDGENANLYMSDDYLVKLVLEF 239
OY 247 HPGFMFTTLYLTGFR-----IVMLTKFDEETFLTKLDYKCSSVILVPTFAILNSEL 302
DB 240 H---IYSLNSVYLCLARVGAAILIKKFDIAQFLLEIPKHVITLPPRPIVLAIAKSP 296
OY 303 LKQYDSNVELASGAPLRSKEIGAVARFPLPGVRGQYGLTETTSAILT-----TPEG 357
DB 297 VNYTDSVRYWASGAPLRSKEIGAVARFPLPGVRGQYGLTETTSAILT-----TPEG 356
OY 358 DDKPGASGVVPLFRKAYIDLTGKTKLGNRRGECVAGKPMKGVNDPNTREIIDE 417
DB 357 DIKSGACGVVNAEMKLYDPDGTGSLPRNGDGLCINGDQIMKYLNDPNTATITKE 416
OY 418 GMLHNGDIDGYDEEKNFIVPLAKSLITYKGVQVPALESVLLQHPNIFDAGVAGVDP 477
DB 417 GMLHNGDIDFIDDELFLVPLKELIKYKGFQVAPAELEALLNHPDISDAAVVPMIDE 476
OY 478 IAGELPGAVVVLKKGKSMTEKEMVDYVAVSOVNAKRLGVRVDVYVGLTKIDGAI 537
DB 478 IAGELPGAVVVLKKGKSMTEKEMVDYVAVSOVNAKRLGVRVDVYVGLTKIDGAI 537

```

```

DB 477 QAGEVPAFVAVNSGTTTEDEVDKFIKQVIFYKRIK-RVFFVETVPKSPSGKILRKDL 535
OY 538 R 538
DB 536 R 536

RESULT 10
ID 4CL3_ARATH STANDARD; PRT; 561 AA.
AC 095777;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE 4-COUMARATE--COA LIGASE 3 (EC 6.2.1.12) (4CL 3) (4-COUMAROYL-COA
DE SYNTHASE 3).
GN 4CL3 OR AT1G65060 OR F16G16.6.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RX MEDLINE=99348176; PubMed=10417722;
RA Ehrling J., Bueltner D., Wang Q., Douglas C.J., Somssich I.E.,
RA Kombirink E.;
RT "Three 4-coumarate:coenzyme A ligases in Arabidopsis thaliana
RT represent two evolutionarily divergent classes in angiosperms.";
RL Plant J. 19:9-20(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.-J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.A., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marzilli A.,
RA Miltscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shin P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820(2000).
RN [3]
RP CATALYTIC ACTIVITY: ATP + 4-COUMARATE + COA = AMP + DIPHOSPHATE +
RP 4-COUMAROYL-COA.
CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF106087; AAD47194.1; -
CC EMBL; AF106088; AAD47195.1; -
CC EMBL; AC009360; AAF06039.1; -
CC HSSP; P08659; ILCT.

```


Db 473 LEALLIGHPDITDAVAVVAAEAGEVPAFVVKSKDSEDDVDQFVSKQVVEFKRT - 531

QY 516 GGYREVDVPEKGLTGKIDGKAIR 538

Db 532 NKVEFTESIPKAPSGKILRKDLR 554

RESULT 15

4CL2_ARATH STANDARD; PRT; 556 AA.

ID 09S725: 09LJ35: 4CL2_ARATH

DT 20-AUG-2001 (Rel. 40, Created)

DT 20-AUG-2001 (Rel. 40, Last sequence update)

DE 4-COMARATE--COA LIGASE 2 (EC 6.2.1.12) (4CL 2) (4-COMAROYL-COA SYNTHASE 2).

GN 4CL2 OR AT3G21240 OR MXL8.10.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

OX NCBI_TaxID=3702;

RA MEDLINE-99348176: PubMed=10417722; Eiling J., Buettner D., Wang Q., Douglas C.J., Somsich I.E., Kombrink E.; "Three 4-comarate:coenzyme A ligases in Arabidopsis thaliana represent two evolutionarily divergent classes in angiosperms."; Plant J. 19:9-20(1999).

RL [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RA MEDLINE-20277480: PubMed=10819329; Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.; "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence features of the regions of 4,504,864 bp covered by sixty P1 and TAC clones."; DNA Res. 7:131-135(2000).

RL DNA Res. 7:131-135(2000).

CC -1- CATALYTIC ACTIVITY: ATP + 4-COMARATE + COA = AMP + DIPHOSPHATE + 4-COMAROYL-COA.

CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.

CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME FAMILY.

CC CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC CC

DR EMBL: AF106085; AAD47192.1; -

DR EMBL: AF106086; AAD47193.1; -

DR EMBL: AB023045; BAB01716.1; -

DR HSSP: P08659; ILCI

DR InterPro: IPR000873; AMP-bind.

DR Pfam: PF00501; AMP-binding.1.

DR PROSITE: PS00455; AMP_BINDING.1.

DR Ligase; Phenylpropanoid metabolism; Multigene family.

KW CONFLICT 247 247 W -> L (IN REF. 2).

FT CONFLICT 265 265 I -> V (IN REF. 2).

FT SEQUENCE 556 AA: 60929 MW: E3C2EAD677529EED CRC64;

Query Match 25.2%; Score 710; DB 1; Length 556;

Best Local Similarity 34.6%; Pred. No. 5,5e-42;

Matches 177; Conservative 92; Mismatches 204; Indels 38; Gaps 8;

Db 57 NGPTEGVYTYADVHTSRKLAAGLNLGKQHDVVMILLPNPEVVLFLASFLGATTT 116

QY 106 PTNEIYTLRELTHSLGISKPTIVFSSKGLKVTIVQK---TWT---AIRTVILDSK 157

Db 117 SANPFTPAEISKOAKASAKLITQSRVYDKIRKIMQNDGVLIYTTSDAIPENCLRFSE 176

QY 158 VDRGYQSMDFIKKNTPOGFKSSFKTYEVNRKQVALINNSGSGTGLPGVQLTHENI 217

Db 177 LTQSEPRVDSIPEKISP-----EDVALPSSSGTGLPGVMLTHKGL 220

QY 218 VTRSHARD---PIYQNVSPGAILITVPFHGFGMT-TLGLTCGRIVMLTKFDE 272

Db 221 VTSVAQVDGENPNLYFNR---DVLICVMPMFHYALNSIMLSLRIGAILIMPFEI 277

QY 273 EFTLTKDQVCSVILPPTLFAILNRELLDKYDLSMLVYIAGGAPLSKEIGENARR 332

Db 278 TLLLEQIQRKAVYVAMVPPVYLAIAKSPETEKYDLSVRVKGGAAPLGKLEDAISAK 337

QY 333 FNLPGVRQGYGLTETTSAILITTPGDDKP-----GASGVVPLFRKAVYIDLTQKTTGPN 387

Db 338 FPNAKLGGYGMTAGPVLAMSLGPAKPPFPVKSAGCTVVRNNAEMKILDPDIDGSLPRN 397

QY 398 RRGECVAGPMLMRGYVNDPEATREIIDEBCWLTGIDIGYDEKHHFIVDRKSLIKYK 447

Db 398 KPGELICINGNIMGYLMDPLATASTIDKGMHTTGVGFTDDDELFIYDRKELIKYK 457

QY 448 GYOVPAPAELESVLLQHPNIPDAGVAGVDPPIAGELPAVVVLKKGKSMTEKEVDVYASQ 507

Db 458 GFQVAPAELESLIGHPINDVAVVANKKEEDAGVPAFVYRKSQNSISEDEIKQVSKQ 517

QY 508 VSNAKRLGCVRFVDEVPKGLTGKIDGKAIR 538

Db 518 VVYFKRI-NKVEFTDSIPKAPSGKILRKDLR 547

Search completed: November 29, 2001, 02:00:07

Job time: 377 sec

